PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Hideki ENDOH, et al.

Docket No: Q82704

Appln. No.: 10/502,279

Group Art Unit: 1639

Confirmation No.: 2122

Examiner: Amber D. Steele

Filed: July 23, 2004

For: METHOD FOR SCREENING A DRUG AMELIORATING INSULIN RESISTANCE

DECLARATION UNDER 37 C.F.R. § 1.132

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

I, Hideki ENDOH, hereby declare and state:

THAT I am a citizen of Japan;

THAT I received the degree of Bachelor of Science from Nagoya University in March 1989, the degree of Master of Science from Kyoto University in March 1991, and the degree of Doctor in Science from Kyoto University in March 1994;

THAT I was employed by Yamanouchi Pharmaceutical Co., Ltd., in 1994 where I held the position of Research Scientist in the Molecular Medicine Research Laboratories, Institute for Drug Discovery Research.

THAT in April 2005, Astellas Pharma, Inc. was formed from the merger of Yamanouchi Pharmaceutical Co., Ltd, and Fujisawa Pharmaceutical Co., Ltd. Since that merger I have been

working as a Research Fellow in the Applied Genomics, Molecular Medicine Research Laboratories, Institute for Drug Discovery Research, Astellas Pharma Inc.

I further declare and state as follows:

I am familiar with the Office Action mailed August 1, 2006, in which the Examiner alleges that the specification does not show definitively that SEQ ID NO:26 contains the promoter region for the FLJ13111 gene, because of the low levels of luciferase activity that are allegedly shown in the specification for the reporter gene construct.

In my expert opinion, the Examiner's position is incorrect.

Specifically, the data in Fig. 9 was incorrectly labeled as representing the ratio of the luciferase activity divided by the β-galactosidase activity. However, in fact, Fig. 9 contains the raw data, i.e., the luciferase activity without being normalized by the β-galactase activity. Nonetheless, Example 14 (and Fig. 9) still prove that SEQ ID NO: 26 has a promoter for FLJ13111.

The reasons are as follows. The method utilized in Example 14 is a method for searching for and identifying a promoter region for a certain gene by comparing the case of a vector which is obtained by inserting the test DNA fragment (DNA fragment to be checked for existence of a promoter region) into pGL3 with the case of pGL3 (a vacant vector containing no DNA fragment) as a control. Such a method itself was well known at the time of filing this application, as is clear from the following literatures (1) to (6), for example. In these literatures, a plasmid obtained by inserting a DNA fragment into pGL3 was prepared and the promoter activity was measured by using, as an index, the reporter activity in the case of the DNA fragment-inserted

plasmid in comparison with the pGL3 (vacant vector). That is, if the reporter activity in the case of the DNA fragment-inserted plasmid is higher than the reporter activity in the case of pGL3 (vacant vector), this means existence of a promoter activity (i.e., the inserted DNA fragment has a promoter region).

It is also common to show the relative activity values (XX folds by taking the value in the case of the vacant vector to be "1") and not to show the raw data values itself or the values obtained by dividing by the transfection efficiency (cf. literatures (2), (4)-(7)).

(1) Masumoto N, Chen J, Sirotnak FM.

Regulation of transcription of the murine gamma-glutamyl hydrolase gene. Delineation of core promoter A and the role of LYF-1, E2F and ETS-1 in determining tumor-specific expression.

Gene. 2002 May 29;291(1-2):169-76.

(cf. page 170, left column, 2,1. Analysis of promoter activity, Fig. 1)

(2) Tzeng SJ, Huang JD.

Transcriptional regulation of the rat Mrp3 promoter in intestine cells.

Biochem Biophys Res Commun. 2002 Feb 22; 291(2):270-7.

(cf. Fig. 2)

(3) Liang L, Major T, Bocan T.

Characterization of the promoter of human extracellular matrix metalloproteinase inducer (EMMPRIN).

Gene. 2002 Jan 9;282(1-2):75-86.

(cf. Fig. 3)

(4) Human Molecular Genetics, 2001, Vol. 10 No. 26, 3101-3109

(5) Biochemical and Biophysical Research Communication 286, 381-387 (2001) (cf. Fig. 2B)

(6) Gene 275 (2001), 93-101

(cf. Fig. 3)

(7) Endocrinology 142(9); 3987-3995, 2001

(cf. Fig. 4)

Fig. 9 shows that the reporter vector having the sequence of SEQ ID NO: 26 showed about 29 times higher luciferase induction in comparison with the control reporter vector pGL3 having no promoter (please compare the lane 1 (left side) and lane 2 (second lane from the left side)). Accordingly, the luciferase activity induced by SEQ ID NO: 26 is not low. The difference between the control reporter vector pGL3 and the reporter vector having the sequence of SEQ ID NO: 26 is significant (the P value is 0.000930527; ** (double stars), when * (single star) is 0.05 or less and ** (double star) is 0.01 or less).

The promoter activity is different among different genes and among different types of cells. It is not common in the field of the reporter assay to use a known reporter for control. As shown in the above cited literatures, it is common to use the value in the case of vacant vector (e.g., pGL3) is taken as the baseline control.

The luciferase activity differs depending on the retaining time of the vector in cells, types of the cell, conditions of the measurement apparatus, etc. The fact that the value differs

depending on the cell type is described in the Promega's web sites (Legend in Fig. 6). As described above, it is quite common to carry out comparison against the case of using the vacant vector under the same experiment. The Examiner compares the values obtained in different experiments in the same way, which is meaningless. For example, it is meaningless to compare the value available on the Promega's web site with the values of this experiment.

Furthermore, when the data in Fig. 9 is divided by the β-galactosidase activity, the results are almost unchanged.

Specifically, the raw data (average value) of the measured luciferase activity (i.e., the values of Fig. 9 at the time of filing date of this application) are 0.0561 (lane 1, vacant vector), 1.6495 (lane 2, a vector having the sequence of SEQ ID NO: 26), 2.5334 (lane 3), 0.1398 (lane 4). The measured value of the β-galactosidase activity as measured in the original experiments, was 0.016 in each of the lanes. Accordingly, the luciferase activity/β-galactosidase activity values are 35.06, 1030.94, 1583.77, and 87.38, respectively. This is shown in Fig. 9', attached hereto.

By dividing with the β-galactosidase activity, the values in the ordinate of Fig. 9' have been changed from those of the original Fig. 9. However, because the measured value of the β-galactosidase activity was the same among the lanes, it is still correct that the value of lane 2 (the case in which a vector having the sequence of SEQ ID NO: 26 was used) was about 29 fold, when the value of the lane 1 (the case of vacant vector) is taken as "1" (1.6495/0.0561 is about 29 and 1030.94/35.06 is about 29).

That is, it is still correct that the value of lane 2 (the case in which a vector having the sequence of SEQ ID NO: 26 was used) was about 29 folds, when the value of the lane 1 (the case of vacant vector) is taken as "1" (1.6495/0.0561 is about 29 and 1030,94/35.06 is about 29).

I declare further that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Date: Nov. 30. 2006

Hideki Endoh,
Hideki ENDOH, Ph.D.

ELSEVIER

Gene 291 (2002) 169–176

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Regulation of transcription of the murine γ-glutamyl hydrolase gene. Delineation of core promoter A and the role of LYF-1, E2F and ETS-1 in determining tumor-specific expression

Naoko Masumoto, Jing Chen, F.M. Sirotnak*

Program in Molecular Pharmacology and Experimental Therapeutics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA

Received 20 November 2001; received in revised form 14 March 2002; accepted 29 March 2002
Received by J.A. Engler

Abstract

Our earlier studies (Gene 268 (2001) 183) showed that transcription of the mouse γ-glutamyl hydrolase (γGH) gene is under the control of two separate promoters widely distributed within the genome. We now report on further studies examining the functional characteristics of the more efficient promoter (promoter A) which is contiguous with an exon 1 (exon A1a) alternate within the main body of the gene. Functional deletion analysis of promoter A in pGL3 transfected in NIH3T3 cells defined a 189 bp region of sequence 26 bp upstream of exon A1a as mediating core promoter transcriptional activity. Further functional deletion analysis and site-directed mutagenesis attributed the activity of this core region to the presence of five SP1 sites within the most upstream 130 bp of sequence and the presence of three *cis*-active elements, Ets-1, Lyf-1 and E2F in the remaining downstream 59 bp of sequence. Electrophoretic gel-mobility shift assays showed that differences in the binding of *trans*-acting factors to these three elements accounts for the tumor-specific expression of this gene. Finally, analysis by fluorescent in situ hybridization revealed that promoter A and the main body of the mouse γGH gene are located between A3 and A5 on chromosome 4. This is an interesting finding in light of our earlier results which located promoter B and two associated alternates of exon 1 of this gene on chromosome 17.

Keywords: Folate polyglutamate hydrolysis; γ-Glutamyl hydrolase gene transcription

1. Introduction

The biosynthesis and turnover of polyglutamates of folates and their analogues has important physiological ramification and pharmacological significance (Kisliuk, 1981; McGuire and Coward, 1984; Moran et al., 1985; Shane, 1989; Barrueco et al., 1992) as determinants of cytotoxicity and therapeutic efficacy of these analogues. Rapid turnover of these polyglutamates occurs through the action of γ-glutamyl hydrolase (γGH, also known as folate polyglutamate hydrolase) after mediated entry of folates and their analogues into lysosomes (Barrueco et al., 1992). Until recently, detailed information on the structure and properties γGH has been lacking. Studies carried out by Yao et al. (1996a) resulted in the purification of γGH from rat hepatoma cells and cloning (Yao et al., 1996a,b; Esaki et al., 1998) of rat, mouse and human complementary DNA's (cDNA's) and, ultimately, the gene in mouse (Esaki et al., 1999) and man (Yin et al., 1999). The mouse gene consists (Esaki et al., 1999; Masumoto et al., 2001) of nine exons with four alternates of exon 1 and is under the control of two promoters, one of which (promoter B) is bi-directional and located in the complement C3 gene locus.

In the current studies, we sought to elucidate some of the properties of the most efficient promotor (promoter A) in regulating transcription of this gene. Our results served to identify a short region of upstream sequence contiguous to exon A1a as responsible for core promoter activity. They also reveal that enhancement of transcriptional activity is associated with the presence of five SP1 sites and the existence of three additional *cis*-elements, Ets-1, Lyf-1 and E2F, within an adjacent 189 bp segment of downstream

Abbreviations: γGH, γ-glutamyl hydrolase; FISH, fluorescent in situ hybridization; PCR, polymerase chain reaction; MEM, minimal essential medium; DTT, dithiotrietol; DAPI, diamidino-2-phenylindol; RNP, ribonuclear protein

* Corresponding author. Tel.: +1-212-639-7952; fax: +1-212-794-4342.
E-mail address: sirotnaf@mskcc.org (F.M. Sirotnak).

0378-1119/02/$ - see front matter
PII: S0378-1119(02)00593-0

sequence. We also provide evidence suggesting that binding of *trans*-acting factors to all three of these elements also determines the marked tumor-specificity for the expression of this gene. Other results obtained by fluorescent in situ hybridization (FISH) showed that the main portion of the mouse γGH gene is located at A3–A5 on chromosome 4.

2. Materials and methods

2.1. Analysis of promoter activity

For these studies, we utilized as a comparator a 1211 bp DNA fragment (Esaki et al., 1999) containing promoter A and exon A1a of the mouse γGH gene inserted into pGL3 basic (Promega, Madison, WI) containing the promoterless luciferase reporter gene. For functional deletion analysis of promoter A, segments of the 1211 bp DNA fragment differing in length were prepared by restriction enzyme digestion or polymerase chain reaction (PCR) and inserted in pGL3 basic. These constructs were then transfected into NIH3t3 cells by the calcium phosphate method (ProfectionR Mammalian Transfection System, Promega). PCH110 (Amersham Pharmacia) expressing β-galactosidase was co-transfected as an internal marker for normalizing the transfection rate. After growth of cells in Eagle's minimal essential medium (MEM) with 10% fetal calf serum for 72 h at 37°C, cell-free extract was prepared and assayed for β-galactosidase and luciferase activities as described by the manufacturer (Promega). Promoter activities were determined relative to pGL3basic alone which was also transfected into NIH3T3 cells. The identity of all of the reporter gene constructs was verified by sequencing.

2.2. Site-directed mutagenesis

These studies were carried out with a Clontech Site-Directed Mutagenesis kit (Clontech). The transcription binding sites Ets-1 (ACCGCAAGCA) and Lyf-1 (CTTCCCAGA), E2F (TTGCGCGC) were identified within promoter A by means of the TFSEARCH (ver. 1.3) program. Mutagenesis primers were designed with 4 or 5 bp mismatches targeting specific *cis*-acting elements within the core promoter sequence. The mutated sequences were AGGTTGAGCA (ETS-1), CTCATAAGA (Lyf-1) and E2F TTGAATAC (E2F). Different selection primers were used depending upon the uniqueness of the encompassed restriction sites. Verification of the mutated segments and the reporter gene constructs was carried out by sequencing (Sequenase 2.0, U.S. BioSciences).

2.3. Polymerase chain reactions

PCR was carried out with the Taq and Tga DNA polymerases (ExpandR Long Template PCR System, Boehringer Mannheim) using the recommended buffer, 300 nM of primers, 1 ng of template DNA and 500 μM dNTP's in a total volume of 50 μl. After an initial denaturation step of 2 min at 94°C, 30 cycles of 10 s at 94°C, 30 s at 55°C and 4 min at 68°C were carried out. The reactions were extended for 7 min at 68°C.

2.4. Electrophoretic mobility shift assay

Nuclear extract were prepared from NIH3T3, Tapper Liver Tumor and P388 cells according to the method of Digman et al. (1983) and stored at −80°C. Protein concentration was determined by the Bradford method (BioRad). Oligonucleotides were synthesized with nucleotide sequence corresponding to each *cis*-acting element of interest. Double-stranded oligonucleotides were generated by annealing complementary nucleotides. The binding reaction was performed at room temp in a vol of 25 μl which contained binding buffer (5% glycerol, 1 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM dithiotrietol, 50 mM NaCl, 10 mM Tris–HCl, pH 7.5), 50 μg/ml poly(dI-dC):poly(dI-dC), 2×10^5 cpm of an α-^{32}P labeled probe and 15 μg nuclear protein. After incubation for 30 min, the samples were electrophoresed on 4% acrylamide gel in 0.5 × TBE buffer. The gel was dried and exposed to X-ray film at −80°C in the presence of an intensifying screen The competition experiments were performed by adding a 200-fold excess of unlabeled probe to the reaction mixture.

The preparation of the oligonucleotide probes were as follows. Double stranded oligonucleotides were generated from *cis* element specific oligonucleotides by annealing with a complementary oligonucleotide at equimolar ratio in ddH_2O at 70°C for15 min and slowly cooling to room temp. These were labeled by filling in the recessed 3′ end with dNTPs including [α-^{32}P]dCTP, using a Klenow fragment of the DNA polymerase. The prepared probes were as follows: E2F, 5′-ctgagcggcgcggccTTGCGCGCcc-3′ annealed to 5′-cgccgcgccggAACGCGCGggagct-3′; Lyf-1, 5′-ctgactgcaCTTCCCAGAggc-3′ annealed to 5′-gacgt-GAAGGGTCTccggact-3′; Ets-1, 5′-agctccgcACCGCAA-GCAcccggac-3′ annealed to5′-ggcgTGGCGTTCGTgggc-ctgacgt-3′.

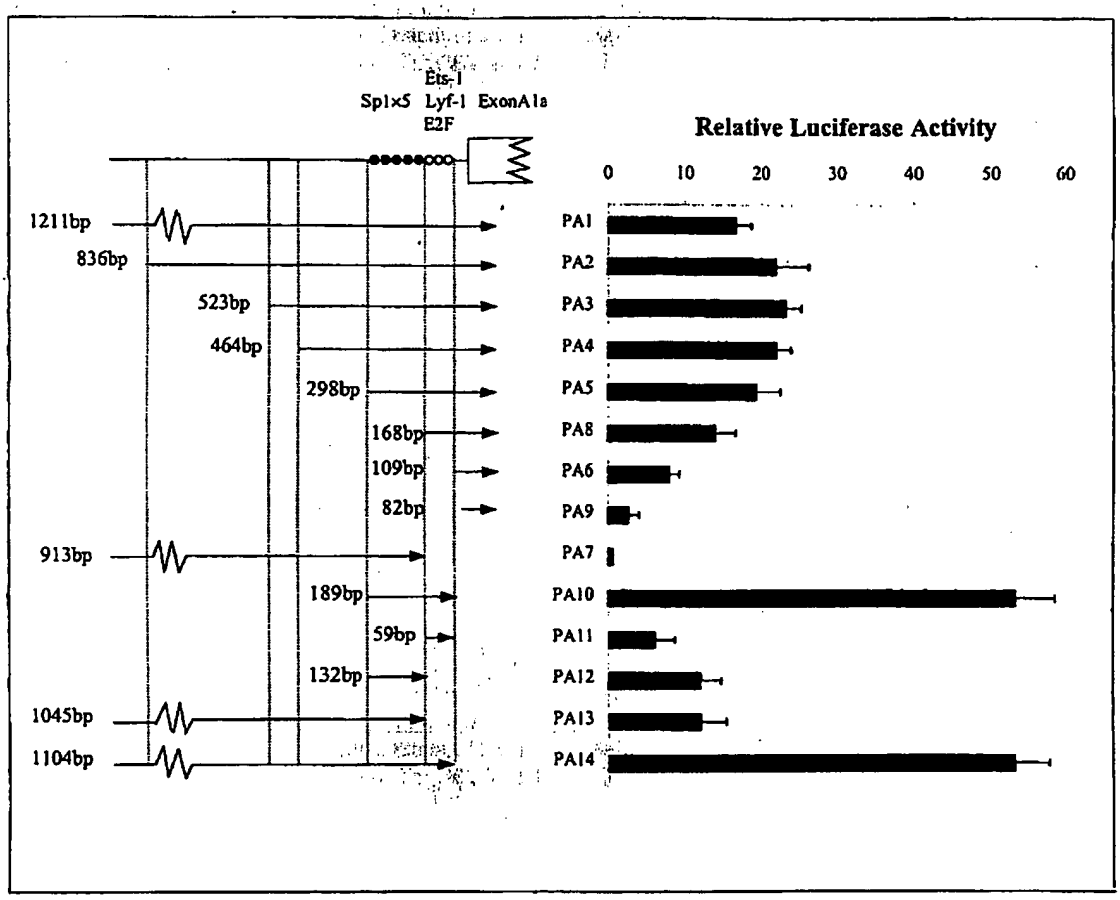
2.5. Cells and culture conditions

NIH3T3, Tapper Liver and P388 cells were obtained from the American Type Culture Collection. These were cultivated in RPMI1640 medium with 10% fetal calf serum.

2.6. FISH mapping

Lymphocytes were isolated from mouse spleen and cultured at 37°C in RPMI1640 medium supplemented with 15% fetal calf serum, 3 μg/ml concanavalin A, 10 μg/ml lipopolysaccharide and 0.5 μM mercaptoethanol. After 44 h the cultured lymphocytes were treated with 0.18 mg/ml BrdU (bromodeoxyuridine) for 14 h. The synchronized cells were washed and recultured at 37°C for 4 h in α-MEM with thymidine (2.5 μg/ml). Chromosome

slides were made by a conventional method as used for human chromosome preparations(hypotonic treatment, fixation and air drying). The γGH DNA probe used for FISH mapping was the γmγGH-4 genomic clone (Esaki et al., 1999) obtained from a λ phage genomic library 15.5 Kb in length and encompassing exons 1–5 and 2.5 Kb of upstream and 3 Kb of downstream sequence This probe was biotinylated with dATP using the GibcoBRL BioNick labeling kit incubating for 1 h at 15°C (Heng et al., 1992). The FISH with amplification was performed according to a published method(Heng et al., 1992; Heng and Tsui, 1993). FISH signals and banding patterns were recorded separately. Images were captured and combined by digital camera and the assignment of the FISH mapping data



PA10 Sequence

```
                                        Sp1                 Sp1
agcttggagcacgctcccagagccttctgtccccggcccgccccggccgccccggcccgc

                  Sp1          Sp1          Sp1
ccccggcccgcctcggcccgcccccggcccgcccccggcccgcccccgcccggcccg

              Ets-1                Lyf-1
cccagctccgcacccgcaagcacccggactgcacttcccagaggcctgagcggcgcggcct

  E2F
tgcgcgccc
```

Fig. 1. Functional deletion analysis of promoter A regulating transcription of the mouse γGH gene. Each construct was made by inserting promoter sequence into pGL3. The data show the luciferase activity obtained with each construct relative to pGL3basic alone. Additional details are provided in the text. The data shown are an average with standard error of the mean of three or more replicate experiments done in duplicate. Statistical analysis was carried out by the Student's *t*-test and the *P* values provided in the text for specific comparisons.

with chromosomal bands was achieved by superimposing FISH signals with the banded chromosomes(Heng and Tsui, 1993).

3. Results

3.1. Functional deletion analysis

We have inserted a 1211 bp 5′upstream region of DNA including 83 bp of exon A1a sequence of the mouse βGH gene into pGL3 and shown in our prior studies (Esaki et al., 1999) that this 1211 bp segment will drive transcription of the luciferase reporter gene in this vector construct. We also inserted into pGL3 different segments of this 5′upstream sequence prepared by restriction enzyme digestion or PCR and determined (Fig. 1) the luciferase activity of these constructs. From the data given in the figure we derived the following information. Although the most upstream 375 bp of this 5′ upstream sequence appears to modestly influence in a positive manner (PA2, PA3 or PA4 versus PA1, $P = <0.05$) transcription in pGL3, overall transcriptional activity is largely influenced by elements within the most downstream 298 bp (PA5). This is also shown (Fig. 1) by the inability of sequence (PA7) upstream of this 298 bp segment alone to sustain transcription (PA7 versus PA1, $P = <0.001$). The functional deletion analysis also revealed that three components important to the activity of PA5 reside in separate but contiguous portions of the 298 bp segment of sequence. This was revealed by a comparison of the activity of PA5–PA8 ($P = <0.05$), PA6 ($P = <0.01$) and PA9 ($P = <0.005$). Interestingly, some promoter activ-

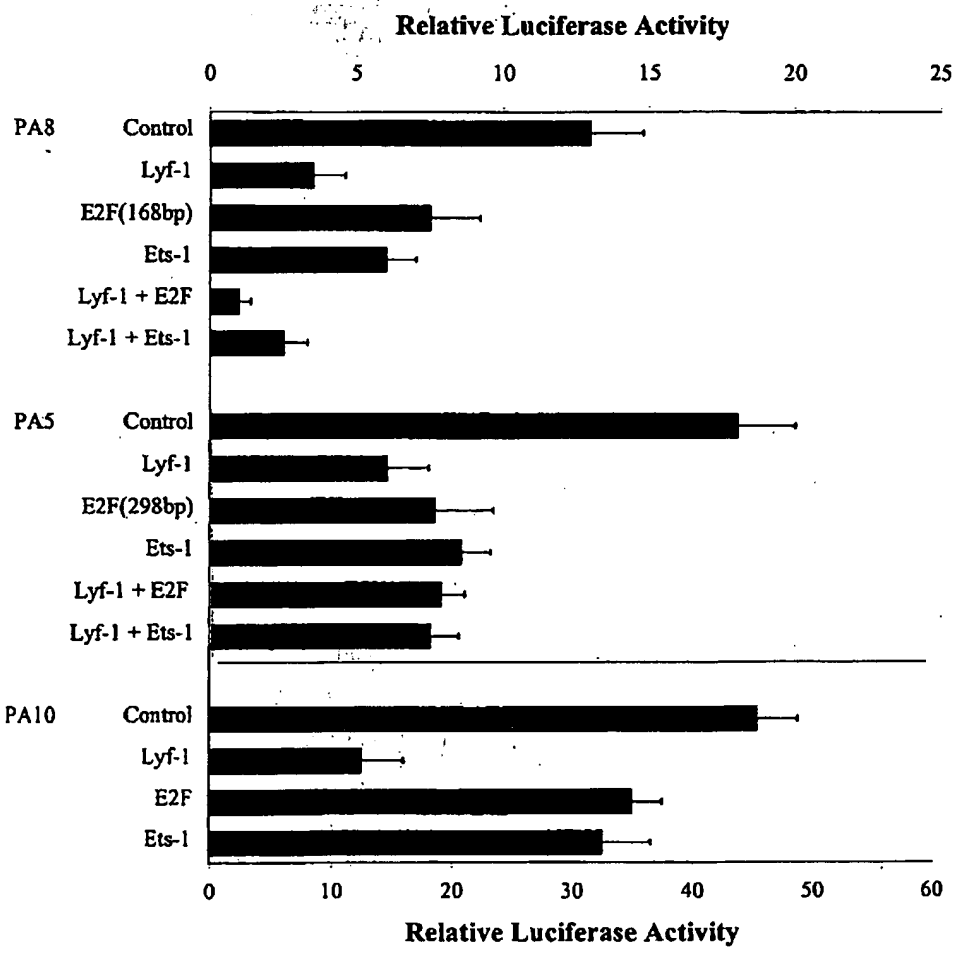


Fig. 2. Site-directed mutagenesis of Lyf-1, E2F and Ets-1 sites within the γGH promoter A. The data show the effect of mutation on relative luciferase activity of pGL3 constructs, PA5, PA8 and PA10. Additional details are given in the text and legend of Fig. 1. The data shown are an average with standard error of the mean of three or more replicate experiments carried out in duplicate. Statistical analysis was carried out by the Student's *t*-test and the *P* values provided in the text for specific comparisons.

ity was sustained by the 82 bp segment alone extending into exon A1a (PA9). However, this 82 bp segment along with the 27 bp of contiguous upstream sequence (PA6) also exhibited suppressor-like activity, since its deletion (PA10 or PA14) from this 298 bp segment (PA5) markedly increased transcription ($P = <0.001$). As the 59 bp downstream portion (PA11) or the 130 bp upstream portion (PA12) of the 189 bp segment incorporated in P10 alone stimulated transcription, albeit to a lessor extent (PA11 and PA12 versus PA10, $P = <0.001$) we conclude that core promoter A activity in NIH3T3 cells reflects the action of *cis*-acting elements in each of these regions of PA10. The 130 bp upstream region of P10 is notable for its content (Fig. 1) of five SP1 sites, while the downstream 59 bp region has putative sites interacting with Ets-1, Lyf-1 and E2F. To what extent these elements would regulate transcription of this gene in other cell types is addressed below.

3.2. Site-directed mutagenesis studies

We used site-directed mutagenesis to evaluate the relative contribution of the Lyf-1, E2F and Ets-1 sites to core promoter activity. These results are shown in Fig. 2. Inactivation of either one of these elements in PA8 resulted in markedly reduced transcription ($P = <0.005$ compared to PA8) with somewhat greater reduction in transcription occurring following inactivation of Lyf-1. The data also show that co-inactivation of Lyf-1 along with either E2F or Ets-1 further reduced transcription($P = <0.01$ for mutation of one versus two elements). Similar results (data not shown) were obtained following co-inactivation of E2F and Ets-1. Some loss of transcription occurred when these same elements were inactivated in PA5 but the effects were muted when compared to the same co-inactivation in PA8. These results might have been anticipated and confirm the role of the multiple SP1 sites in PA5 on transcription of this gene.

3.3. Electrophoretic gel-mobility shift analyzes

Northern blot analysis has shown (Esaki et al., 1999) that the γGH gene is differentially expressed among mouse tumors. Some tumors express high levels of this gene while others express low levels. We carried out a series of mobility shift assays to determine to what extent cis elements that bind Lyf-1, E2F and Ets-1 might play a role in determining this tumor-specificity for expression of this mouse gene. The results obtained with nuclear extracts of NIH3T3 and Tapper Liver Tumor cells, both of which express high levels of βGH, and P388 cells expressing low levels of this gene are shown in Figs. 3 and 4. These results document (Fig. 3) a marked shift in the mobility of the Ets-1, Lyf-1 and E2F specific probes in the presence of

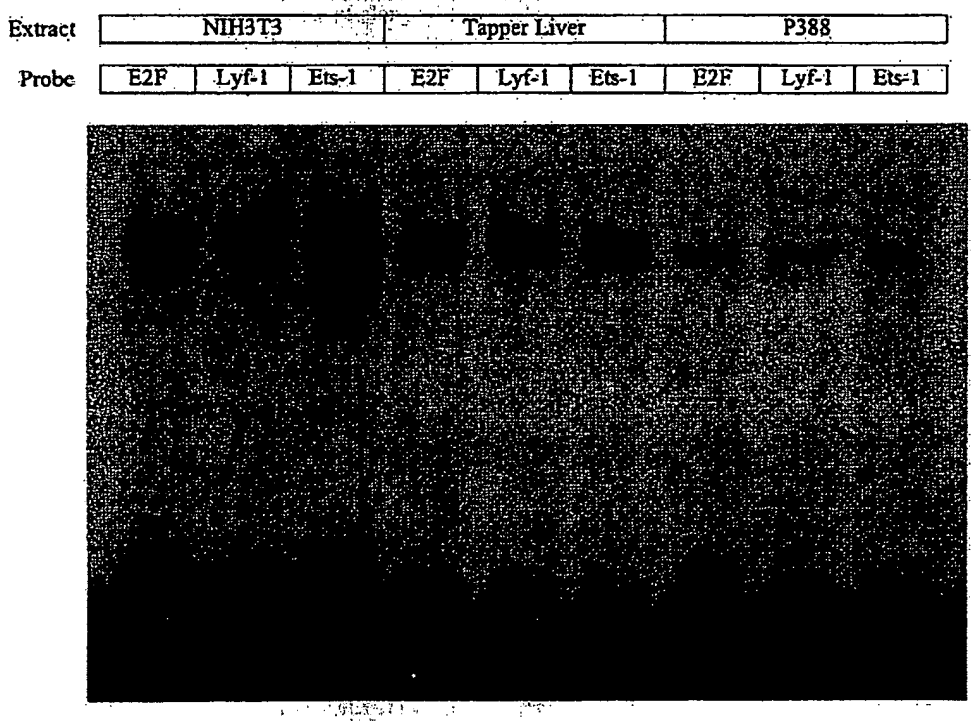


Fig. 3. Electrophoretic gel-mobility shift determination of the interaction between Lyf-1, E2F and Ets-1 elements within the γGH promoter and nuclear protein from NIH3T3, Tapper Liver Tumor and P388 cells. See text for additional details. Double stranded oligonucleotides 20–25 bp in length were generated from cis element specific oligonucleotides by annealing with complementary oligonucleotides and labeled by filling in the 3′ recessed end with dNTP's including [α-^{32}P]CTP using a Klenow fragment of DNA Polymerase. See text for additional details. One of several representative experiments.

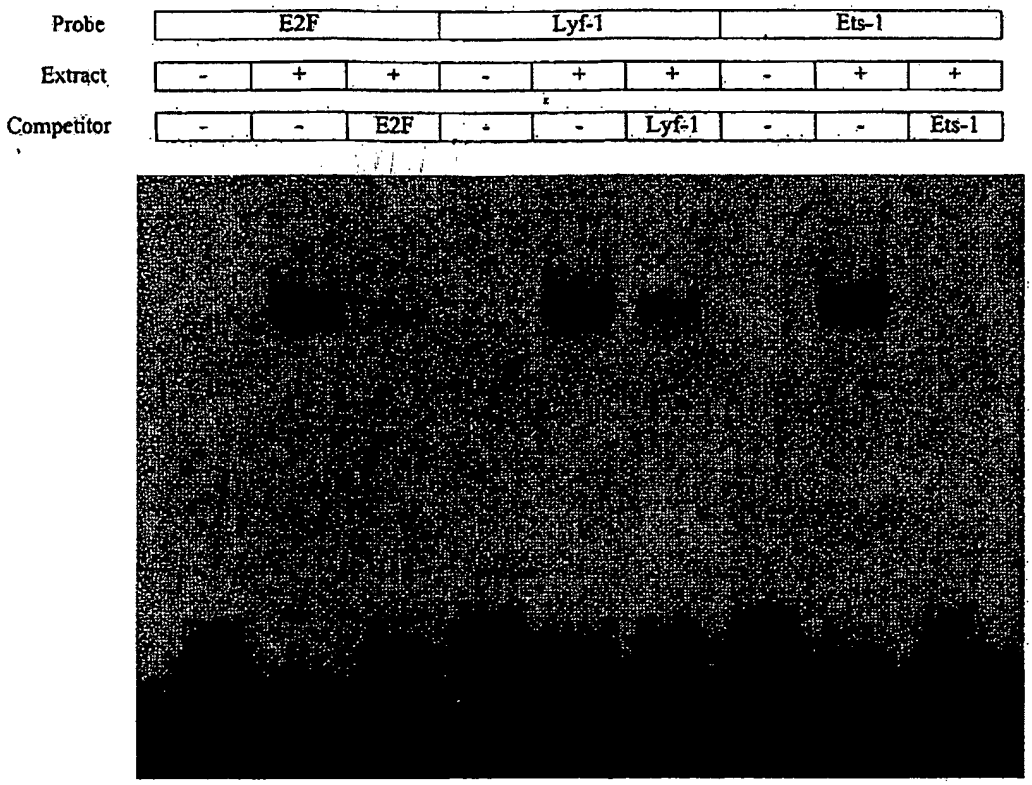


Fig. 4. Electrophoretic gel-mobility shift determination of the interaction between Lyf-1, E2F and Ets-1 elements within the γGH promoter and nuclear protein from Tapper Liver Tumor in the presence and absence of competing oligonucleotides. See text and legend of Fig. 3 for additional details. One of several replicate experiments.

nuclear extract from all three cell types. However, the level of probe shifted in the case of nuclear extract from NIH3T3 and Tapper Liver tumor cells was substantially greater than with extract from P388 cells. Other data in Fig. 4 show that the gel shift with the three probes with nuclear extract from Tapper Liver Tumor cells was markedly suppressed when competing non-radioactive probe was added to the reaction mixture. Similar results (data not shown) with these competing probes were also obtained with nuclear extract from the other tumors.

3.4. FISH studies

Using a γGH gene specific probe, we carried out FISH on lymphocytes from mouse spleen. Under the conditions employed, the detection efficiency for a single chromosome was 78% for a total of 100 mitotic figures. The DAP1 banding was performed (Fig. 5) to identify chromosome 4 on which the hybridization signal occurred (Fig. 6). The gamma GH gene was mapped to the A3–A5 region of this chromosome. A FISH signal was not detected on any chromosome other than chromosome 4.

4. Discussion

The results presented here show that core promoter A activity regulating transcription of the mouse γGH gene is attributable to a 189 bp segment of sequence in the 5′ upstream region of this gene. Activity of this core promoter appears to rely on the presence of five SP1 sites in an upstream segment of 130 bp and to the presence of Lyf-1, E2F and Ets-1 elements further downstream. Site-directed mutagenesis showed that the contribution made by this Lyf-1 site may be greater than the other two elements and that the role of these three elements is strongly influenced by the SP1 sites within the adjacent upstream region. Interesting as well, were our findings suggesting that the marked tumor-specificity for expression of the γGH gene related to differences in the level of *trans*-acting factors binding at these three sites in the various tumor types studies. High levels of expression of the γGH gene found (Esaki et al., 1998) in Tapper Liver Tumor cells compared to P388 cells can now be explained on this basis. E2F, an Rb pathway component, and Ets-1 both of which are well described (Heinemeyer et al., 1998) transcription factors found to be involved in the regulation of a number of genes. Lyf-1 was first identified

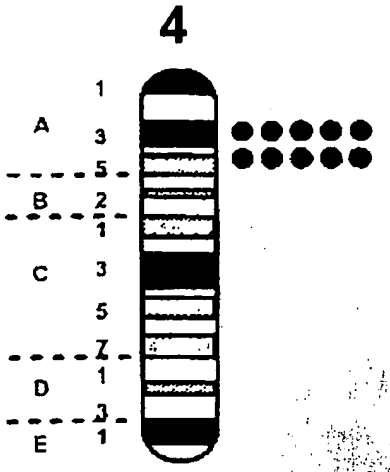


Fig. 5. Diagram of FISH mapping results for the mouse γGH probe. Each dot in the figure represents the double FISH signals detected on mouse chromosome 4. See text for additional details. Fish signals and the diamidino-2-phenylindol (DAPI) banding pattern were recorded separately. Images were captured and combined by digital camera and the assignment of the FISH mapping data with chromosomal bands was achieved by superimposing FISH signals with DAPI banded chromosomes(Heng and Tsui, 1993). One of several determinations.

(Lo et al., 1991) as regulating transcription of the TdT gene in lymphocytes and is especially elevated during B and T lymphocyte development. There is also evidence (Ernst et al., 1993) that Lyf-1 activates transcription of TdT through its interaction with Ets-1. Therefore, it was of interest to find that both *cis*-acting elements have a role in the present context and were found in close proximity to each other within the core promotor A region of the γGH gene.

Our earlier studies have shown (Masumoto et al., 2001) that promoter B and Exons B1a and B1b of the mouse γGH

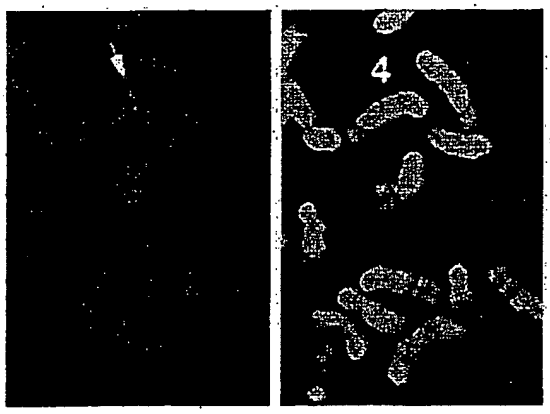


Fig. 6. A representative example of FISH mapping with the mouse γGH probe. The left panel shows the FISH signal. The right panel shows the same mitotic image stained with DAPI to identify mouse chromosome 4. See the text and the legend of Fig. 5 for additional experimental details.

gene are located on chromosome 17 within the complementary DNA strand of the complement C3 gene. They also showed that promoter B was bi-directional, regulating transcription of a second, but unknown, gene in addition to the γGH gene. These earlier studies showed that promoter B and the associated alternates of exon 1 were substantially distal (>45 Kb) from the main body of the γGH gene. Our findings from the FISH analysis described here now show that the main body of this gene is located on chromosome 4. This is an interesting result showing that promoter B residing on chromosome 17 can regulate transcription of the major portion of this gene which is located on chromosome 4.

The underlying cytogenetic basis for the results of the FISH analysis is not known but most likely represents a translocation event. However, the possibility that the presence of promoter B and exons B1a and B1b on chromosome 17 resulted from translocation of a duplicated segment can be ruled out by the fact that a FISH signal on any other chromosome but chromosome 4 was not detected. The existence of this sequence on chromosome 4 as a pseudogene can be ruled out for the same reason and the fact that this portion of the γGH gene on chromosome 17 is transcribed as constituted in variant II messenger RNA (mRNA) (Esaki et al., 1998).

Although the interchromosomal location of promoter B and exons B1a and B1b compared to the main body of the mouse γGH gene is an unusual cytogenetic finding, the ability of promoter B to regulate transcription of this gene is not inconsistent with current dogma (Dreyfuss et al., 1993). Promoter B initiates transcription of exons B1a and B1b and the rest of the gene through splicing of exon B1b to exon 2 as in variant II mRNA (Esaki et al., 1998). The transcript incorporating exons B1a and B1b along with all other precursor mRNA's transcribed from other genes including γGH on chromosome 4 enter the same pool in the nucleus. Specificity of splicing within this precursor RNA pool relies on binding of ribonuclear protein (RNP) proteins within the spliceosome to each donor/acceptor site (Black, 1995). Therefore, the chromosomal origin of the precursor mRNA is irrelevant as long as it is able to enter the global precursor mRNA pool where binding of the appropriate RNP proteins guides the correct splicing events.

5. Conclusions

Core promoter A regulating transcription of the γGH gene is comprised of a 130 bp segment of 5′upstream sequence encoding multiple SP1 sites contiguous to a 59 bp downstream segment of sequence encoding three additional *cis*-acting elements, Ets-1, Lyf-1 and E2F. The interaction of *trans*-acting factors with these three elements plays a role in determining tumor-specific expression of this gene. Promoter A and the main body of the γGH gene are located at A3–A5 on chromosome 4. This was in sharp contrast to our earlier findings which showed that

promoter B and two associated alternates of exon 1 were located on chromosome 17. The cytogenetic basis for the presence of this promoter and the associated alternates of exon 1 on chromosome 17 is unknown, but through splicing of the precursor mRNA's transcribed from this chromosomal locus and the γGH locus on chromosome 4, variant II γGH mRNA (Esaki et al., 1998) must be constituted.

Acknowledgements

This work was supported in part by grants CA08748 and CA56517 from the National Cancer Institute.

References

Barrueco, J.R., O'Leary, D.F., Sirotnak, F.M., 1992. Metabolic turnover of methotrexate polyglutamates in lysosomes derived from S180 cells. Definition of a two-step process limited by mediated lysosomal permeation of polyglutamates and activating sulfhydryl compounds. J. Biol. Chem. 267, 15356–15361.

Black, D.L., 1995. Finding splice sites within a wilderness of RNA. RNA 8, 763–771.

Digman, J.D., Lebovitz, R.M., Roeder, R.G., 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acid Res. 11, 1475–1489.

Dreyfuss, G., Matunis, M.J., Pinol-Roma, S., Burd, C.G., 1993. hnRNP proteins and the biogenesis of mRNA. Ann. Rev. Biochem. 62, 289–321.

Ernst, P., Hahm, K., Smale, S.T., 1993. Both Lyf-1 and an Ets protein interact with a critical promoter element in the murine terminal transferase gene. Mol. Cell. Biol. 13, 2982–2992.

Esaki, T., Roy, K., Yao, R., Galivan, J., Sirotnak, F.M., 1998. Cloning of mouse γ-glutamyl hydrolase in the form of two cDNA variants with different 5′ends and encoding alternate leader peptide sequences. Gene 219, 37–44.

Esaki, T., Masumoto, N., Hayes, P., Chen, J., Sirotnak, F.M., 1999. Organization and structure of the mouse γ-glutamyl hydrolase gene and the functional demonstration of its promoter. Gene 234, 93–100.

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, O.V., Ignatieva, E.V., Ananko, E.A., Podkolodnaya, O.A., Kolpakov, F.A., Podkolodny, N.L., Kolchanov, N.A., 1998. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucleic Acids Res. 26, 362–367.

Heng, H.H.Q., Tsui, L.-C., 1993. Modes of DAPI banding and simultaneous in situ hybridization. Chromosoma 102, 325–332.

Heng, H.H.O., Squire, J., Tsui, L.-C., 1992. High resolution mapping of mammalian genes by in situ hybridization to free chromatin. Proc. Natl. Acad. Sci. USA 89, 9509–9513.

Kisliuk, R.L., 1981. Pteroylpolyglutamates. Mol Cell. Biochem. 39, 331–345.

Lo, K., Landau, N.R., Smale, S.T., 1991. Lyf-1, a transcriptional regulator that interact with a novel class of promoters for lymphocyte-specific genes. Mol. Cell. Biol. 11, 5229–5243.

Masumoto, N., Isaki, T., Sirotnak, F.M., 2001. Additional organizational features of the murine γ-glutamyl hydrolase gene. Two remotely situated exons within the complement C3 gene locus encode an alternate 5′ end and proximate ORF under the control of a bidirectional promoter. Gene 268, 183–194.

McGuire, J.J., Coward, J.K., 1984. Folates and pterins. In: Blakely, R.L., Benkovic, S.J. (Eds.). Pteroylpolyglutamates: Biosynthesis, Degradation and Function, 1. Wiley Interscience, New York, pp. 135–190.

Moran, R.G., Colman, P.D., Rosowsky, A., Forsch, R.A., Chan, K.K., 1985. Structural features of 4-amino antifolates required for substrate activity with mammalian folylpolyglutamate synthetase. Mol. Pharmacol. 27, 155–166.

Shane, B., 1989. Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism. Vitam. Horm. 45, 263–335.

Yao, R., Nimec, Z., Ryan, T.J., Galivan, J., 1996a. Identification, cloning and sequencing of a cDNA coding for rat γ-glutamyl hydrolase. J. Biol. Chem. 271, 8525–8528.

Yao, R., Schneider, E., Ryan, T.J., Galivan, J., 1996b. Human γ-glutamyl hydrolase: cloning and characterization of the enzyme expressed in vitro. Proc. Natl. Acad. Sci. USA 93, 10134–10138.

Yin, D., Chane, K.J., Maculso, C.R., Galivan, J., Yao, R., 1999. Structural organization of the human γ-glutamyl hydrolase gene. Gene 238, 463–470.

Biochemical and Biophysical Research Communications 291, 270–277 (2002)
doi:10.1006/bbrc.2002.6442, available online at http://www.idealibrary.com on IDEAL®

Transcriptional Regulation of the Rat Mrp3 Promoter in Intestine Cells

Shwu-Jen Tzeng and Jin-ding Huang[1]
Department of Pharmacology, College of Medicine, National Cheng Kung University, Tainan 70101, Taiwan, Republic of China

Received January 9, 2002

The promoter of rat multidrug resistance protein 3 (*Mrp3*) has been cloned and analyzed in the rat intestinal cell line (IEC-18 cells). A series of 5′ deletion mutants of the *Mrp3* promoter region were constructed and placed into the pGL3-Basic vector (luciferase reporter gene). Deletion analysis of the *Mrp3* promoter identified a basal transcription element at −123/−106, two negative response regions at −2723/−1128 and −530/−443, respectively, and two positive response regions at −1063/−943 and −302/−157. Further site-directed mutagenesis analysis and gel mobility shift assays provided evidence for Sp1 and Sp3 binding within −123/−105 regions. These studies indicate that Sp1 and Sp3 may be involved in the regulation of the rat Mrp3 gene. © 2002 Elsevier Science (USA)

Key Words: **rat multidrug resistance protein 3 (Mrp3); organic anion transporter; promoter analysis; intestinal cells; Sp1; Sp3.**

Members of the multidrug resistance protein (MRP) family mediate unidirectional ATP-dependent transport of anionic conjugates and amphiphilic anions. Seven members (*MRP1-7*) have been identified (1). Their tissue distribution and cellular localization are different. Substrates for MRP3 include glucuronosyl and sulfated conjugates, whereas glutathione and glutathione conjugates are relatively poor substrates for MRP3 when compared with MRP1 and MRP2 (2). In addition, expression of MRP3 in human embryonic kidney 293 (HEK 293) or ovarian carcinoma cells results in resistance to some anticancer agents, e.g., etoposide, vincristine, and methotrexate (3–5). These data suggest the possible function of MRP3 in physiology and its potential contribution to drug resistance of cancer cells.

Rat Mrp3 is predominantly expressed in intestines (6). Although the primary function of the small intestine is to absorb food and water, it also serves as a major portal of entry for many chemicals, including drugs and toxic compounds in the environment. It therefore is one of the greatest exposures to xenobiotics in the body. The epithelial cells of the small intestine are able to catalyze numerous biotransformation reactions and provide the first site for metabolism of orally ingested xenobiotics. Numerous enzymes catalyzing phase I reaction, e.g., cytochrome P-450, and phase II reactions, e.g., UDP-glucuronosyltransferasse, glutathione *S*-transferases, and sulfotransferases, have been localized to enterocytes (7, 8). Conjugation with glucuronic acid, glutathione, and sulfate represents the major phase II pathways identified in the small intestine. The relevance of the intestine in conjugating reactions is not restricted to xenobiotics. Several endogenous compounds are also efficiently metabolized by this tissue. Bilirubin and steroid hormones are the most common endogenous substances that undergo intestinal conjugation (9, 10). Transport of substrates into the intestinal cells and/or release of these conjugated metabolites rather than the biotransformation enzyme activity have been postulated to be the rate-limiting steps in overall intestinal metabolism (11, 12). MRP3 is also extensively expressed on the basolateral membrane of human cholangiocytes and accepts bile salts as substrates, suggested that MRP3/Mrp3 plays a significant role in the enterohepatic circulation of bile salts (13, 14). Therefore, it is important to know the regulation mechanism of Mrp3 gene expression in the intestine.

MATERIALS AND METHODS

Cell culture. Rat intestine cell line (IEC-18) was obtained from Culture Collection and Research Center, Food Industry and Development Institute (CCRC 60230). The IEC-18 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 5% bo-

[1] To whom correspondence and reprint requests should be addressed at Department of Pharmacology, National Cheng Kung University, Medical Center, 1 University Road, Tainan 70101, Taiwan, Republic of China. Fax: 011-886-6-2749296. E-mail: jinding@mail.ncku.edu.tw.

0006-291X/02 $35.00

TABLE 1

The Oligonucleotide Sequence of Primers

Name	Sequence[a]	Strand	Location[f]
GSP 1[a]	5′-AGCGGTCCATGTGGC-3′	Antisense	−5 ~ +10
GSP 2[a]	5′-TGGTCTTGGGATCTCAGTTCAG-3′	Antisense	−19 ~ −31
−1128F/*Sac*I[b]	5′-ataGAGCTCCTAGTGACTCTGATAAACTCC-3′	Sense	−1128 ~ −1106
−1063F/*Sac*I[b]	5′-ataGAGCTCTAGTACGACGCTCCTC-3′	Sense	−1063 ~ −1047
−943F/*Sac*I[b]	5′-ataGAGCTCGAGAGCAGCCTGACAAATTC-3′	Sense	−943 ~ −923
−873F/*Sac*I[b]	5′-ataGAGCTCGAAGCACAGACAGGCAG-3′	Sense	−873 ~ −856
−806F/*Sac*I[b]	5′-ataGAGCTCCAAGGCTACACAGAGAAAC-3′	Sense	−806 ~ −787
−680F/*Sac*I[b]	5′-gtaGAGCTCTCTACATGCAATACGCACGTG-3′	Sense	−680 ~ −659
−573F/*Sac*I[b]	5′-ataGAGCTCGAAGAGAGCGCAGTG-3′	Sense	−573 ~ −557
−530F/*Sac*I[b]	5′-ataGAGCTCGAAGGCAGCAACTTAAGG-3′	Sense	−530 ~ −512
−443F/*Sac*I[c]	5′-ataGAGCTCACCTGCGCAGAGACACGTCAC-3′	Sense	−443 ~ −422
−302F/*Sac*I[c]	5′-ataGAGCTCAGGACTTCCTTTTACAGTGAG-3′	Sense	−302 ~ −282
−157F/*Kpn*I[c]	5′-ata**GGTACC**CTGTGTTGCAATCGCTGGGACTG-3′	Sense	−157 ~ −134
−88F/*Kpn*I[c]	5′-ata**GGTACC**AAGGAGTGCCTTCGGAGCTG-3′	Sense	−88 ~ −68
MS-F1/*Kpn*I[c]	5′-ata**GGTACC**ACCTTTGCAGAGACACGTCAC-3′	Sense	−443 ~ −422
MS-F2[d]	5′-CAACTACCTTCCTTACTCTGC-3′	Sense	−215 ~ −193
MS-R2[d]	5′-GCAGAGTAAGGAAAGGTAGTTGACAG-3′	Antisense	−219 ~ −193
MS-F3[d]	5′-TGACCCAGAGGTTTGGGCGGGAAG-3′	Sense	−109 ~ −86
MS-R3[d]	5′-CTTCCCGCCCAAACCTCTGGGTCA-3′	Antisense	−109 ~ −86
MS-F4[d]	5′-TGACCCAGAGGGCGGGTTTGGAAGGAG-3′	Sense	−109 ~ −83
MS-R4[d]	5′-CTCCTTCCAAACCCGCCCTCTGGGTCA-3′	Antisense	−109 ~ −83
MS-F5[d]	5′-TGGAGAGAGTTGGTTCCG-3′	Sense	−63 ~ −46
MS-R5[d]	5′-GCCCCGGAACCAACTCTCTC-3′	Antisense	−61 ~ −42
MS-F3, 4[d]	5′-TGACCCAGAGGTTTGGTTTGGAAGGAG-3′	Sense	−109 ~ −83
MS-R3, 4[d]	5′-CTCCTTCCAAACCAAACCTCTGGGTCA-3′	Antisense	−109 ~ −83
−19R/*Xho*I[c]	5′-ata***CTCGAG***TGGTCTTGGGATCTCAG-3′	Antisense	−36 ~ −19

[a] Primers used for Genome Walker library screening and 5′-RACE assay.
[b] Primers used to obtain the 5′-deletion construct for reporter gene analysis.
[c] Primers used to obtain the 5′-deletion and site-specific mutagenesis constructs for reporter gene analysis.
[d] Primers used for amplification of the site-specific mutagenesis fragments.
[e] Lowercase denotes nucleotides flanking the restriction site for improving cleavage close to the end of DNA fragments. *Sac*I, *Kpn*I, and *Xho*I sites are indicated as underlined, bold, and italic–bold letters, respectively. Double-underlined letters denote nucleotides of site-specific mutagenesis.
[f] Nucleotide positions relative to the transcription start site (+1) of the rat Mrp3 gene.

vine fetal serum, and 0.1 U/ml insulin (Gibco BRL, Grand Island, NY) in 5% CO_2 at 37°C.

Cloning of the 5′-flanking region of the rat Mrp3 gene. The 5′-flanking region of the rat *Mrp3* gene was isolated by Genome Walker kits (Clontech, Palo Alto, CA) following the protocol of the manufacturer. The primers used were listed in Table 1. The PCR product was ligated with pCRII vector and sequenced by an Applied Biosystems Model 377A (Applied Biosystems, Foster City, CA).

Construct of reporter plasmids. Deletion fragments of the *Mrp3* promoter were generated by PCR using primers containing *Kpn*I or *Sac*I and *Xho*I restriction sites. The 2723-bp *Sac*I–*Xho*I DNA fragment was cloned directly upstream of the luciferase gene (pGL3-Basic plasmid; Promega, Madison, WI) to generate pWT-2723. A clone containing the base −443 DNA fragment was used as template to amplify base −302, −157, and −106 region upstream of the ATG initiation codon. A series of 5′ deletion fragments of the rat *Mrp3* 5′-flanking region, each differing in the size approximately 30~120 bp (Fig. 2A), were amplified by PCR using the 1198 bp promoter region of rat *Mrp3* cloned into pGL3-Basic plasmid as a template. PCR fragments were generated using different forward primers (−1128F/*Sac*I, −1063F/*Sac*I, −943F/*Sac*I, −873F/*Sac*I, −806F/*Sac*I, −680F/*Sac*I, −573F/*Sac*I, and −530F/*Sac*I) and the same reverse primer (−19R/*Xho*I). Therefore, all fragments varied in their 5′ ends, but all their 3′ ends terminated 19 nucleotides upstream from the translation initiation site. All oligonucleotide primers used in this study are listed in Table 1. PCR products were digested with *Sac*I and *Xho*I to create the cohesive ends, purified and then ligated into the *Sac*I/*Xho*I site upstream from the luciferase gene in the pGL3-Basic vector. All the deletion constructs were confirmed by DNA sequencing. Plasmid-preparation purity was confirmed by an $A_{260/280}$ value of >1.6, and supercoiling of DNA was established by the appearance of a quickly migrating band on agarose gel electrophoresis prior to use in transfection experiments.

Mutant constructs were prepared by replacing the wild-type sequences with mutated sequences generated by the PCR amplification using the constructs pWT-443, pWT-302, and pWT-157 as templates, respectively. The primer designs are listed in Table 1. The PCR-based procedures of mutagenesis were an application of "overlap extension." The first round of PCR is performed using mutant primer (MS-F3 or MS-R3) prerun 10 cycles, then added the flanking primer (−19R/*Xho*I or −157F/*Sac*I) for another 30 cycles. The PCR product with site-directed mutagenesis was digested with optimal restriction enzymes and directly inserted into the *Sac*I/*Xho*I or *Kpn*I/*Xho*I site of upstream of pGL3-Basic vector.

Transfection and dual-luciferase assay. Transient transfection was performed using lipofectAMINE. IEC-18 cells (approximately 4×10^5 cells) were placed onto a six-well plates and were cotransfected with a reporter plasmid (0.3 μg) and the pRL-TK vector (0.3 μg) (Promega, Madison, WI) as an internal control. Five hour after the DNA transfection, cells were replaced with fresh complete me-

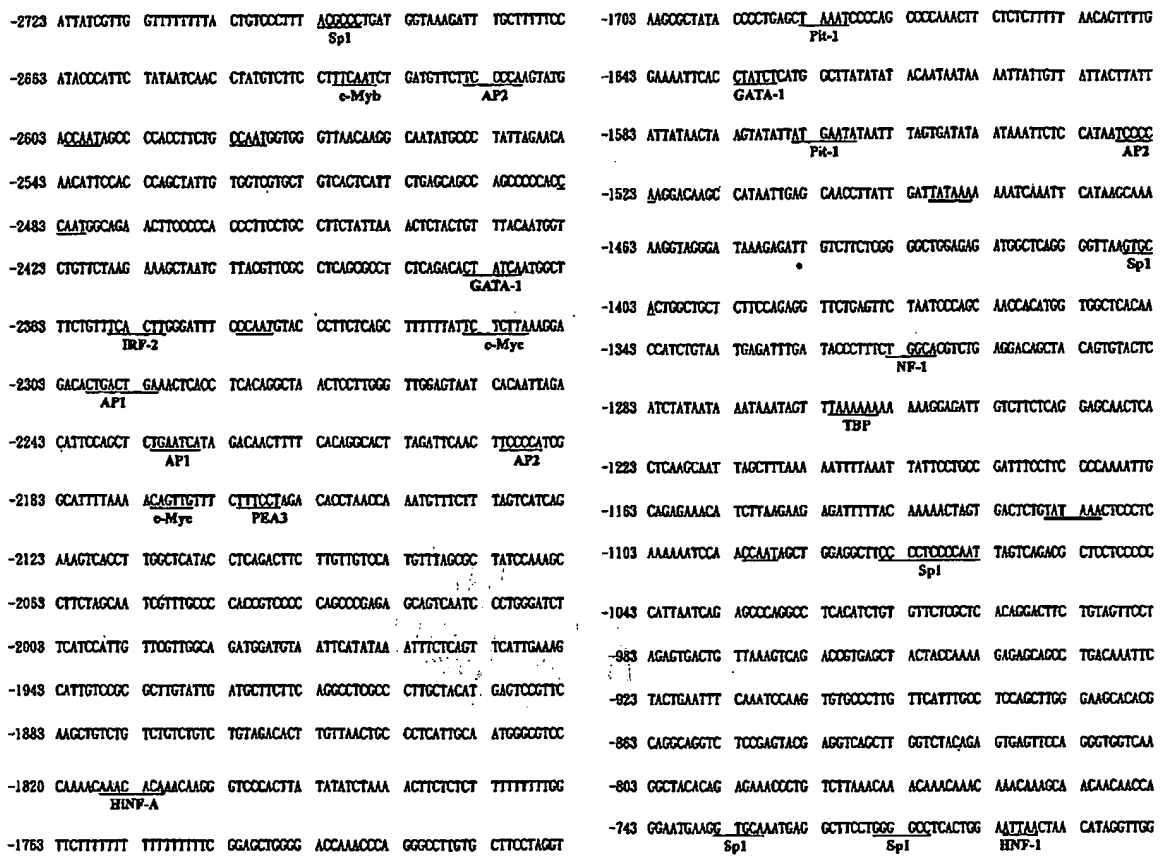


FIG. 1. Nucleotide sequence of proximal 5′-flanking, partial coding and intron 1 of the rat *Mrp3* gene (GenBank Accession No. Y039030). Number of nucleotides is relative to the translation start site. The coding region with standard single-letter amino acid abbreviations is displayed underneath the respective codon. Putative transcriptional regulatory elements are underlined and indicated in boldface type. The putative TATA box is double-underlined.

dium. After 72 h, the cells were washed with phosphate-buffered saline, followed by a dual-luciferase assay according to the manufacturer's instructions (Promega).

Preparation of nuclear extracts. Nuclear extracts were prepared from IEC-18 and Drosophila SL2 cells according to the method of Chen and Chang (15) with minor modifications.

Electrophoretic mobility shift assays. The sequence of wild type and mutant oligonucleotides used as probes and competitors is shown at the Figure 3A. Oligonucleotides were synthesized, annealed to form double strands, and labeled with [γ-^{32}P]ATP by T4 polynucleotide kinase. Reactions were carried out by incubating 6–9 μg of nuclear extract in 1× gel shift buffer [4% glycerol, 1 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT, 50 mM NaCl, 10 mM Tris–HCl (pH 7.5), and 0.05 mg/ml poly(dI–dC) · poly(dI–dC)] for 10 min at room temperature. Probes (20,000 cpm) were added to the mixtures to a final volume of 10–15 μl and incubated for another 20 min at room temperature. Antibodies to Sp1 and Sp3 were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Antibodies were added to nuclear extracts for 20 min at room temperature prior to the addition of labeled WT (wild type) probe. DNA–protein complexes formed were then analyzed by electrophoresis on nondenaturing 4% polyacrylamide gels containing 0.5× TBE buffer at 350 V at cold room until the bromophenol blue dye is three-fourths of the way down the gel. The gel was dried and exposed by MR-X ray film (Amersham) at −70°C with an intensify screen. Oligonucleotides used as probes as follows: WT, 5′-ACCCAGAGGGCGGG GCGGGAAG-3′; MS (mismatched type), 5′-ACCCAGAGG*TTT*GGGCGGGAAG-3′; Sp1, 5′-ATTCGATG GGGGCGGGGCGAGC-3′.

RESULTS AND DISCUSSION

The 5′-flanking region from −2723 to +45 of the rat *Mrp3* gene was isolated by a genomic walking method. The 3′-sequence was identical with the published rat *Mrp3* cDNA sequence begins at nt −48 to +45. Comparison of the 5′-flanking region to sequences available in both GenBank and EMBL databases did not reveal any matches. This sequence was extensively searched for homology to previously described regulatory elements (shown in Fig. 1). Analysis of the 5′-flanking sequence revealed the presence of various consensus transcription factor elements using WWW Signal Scan IMD Search Service (http://bimas.dcrt.nih.gov/molbio/matrixs) and the database program TFSERCH program (http://pdap1.trc.rwcp.or.jp/research/db/TFSEARCH.

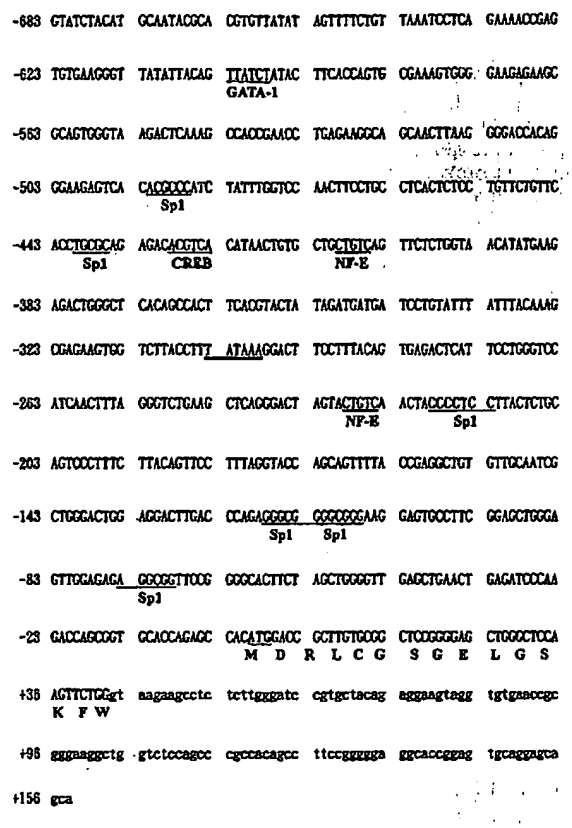


FIG. 1—*Continued*

html and http://transfac.gdf.de/pat Search/) (16). There were three TATA boxes and many GC boxes know as binding sites for transcription factor Sp1 adjacent to the translation start site.

To identify the sequences with the promoter activity in the 5′-flanking region of the *Mrp3* gene, various truncated 5′-flanking sequences fused to the luciferase reporter gene were introduced into rat IEC-18 cells by transfection (Fig. 2). Luciferase activity was normalized for transfecting using a *Renilla* luciferase expression plasmid as an internal control. As can be seen in Fig. 2A, the fusion of *Mrp3* promoter sequences from −2723 to −19 upstream of the luciferase (LUC) gene resulted in the 39-fold induction of LUC activity when compared with pGL3-Basic in IEC-18 cells. The luciferase activity in IEC-18 cells increased in progressive 5′-deletion constructs until the deletion reached the base −1128, suggesting the presence of multiple repressive sequences. A further deletion to base −943 reduced its activity by 68%. A successive deletion to base −530 had little effect on luciferase activity. Deletion from base −530 to −433 increased it activity by 2.2-fold. It was suggested there were two regions of a prominent repressive effect on reporter expression between bases −2723/−1128, and −530/−443. The pWT-443 and pWT-302 constructs showed about 50-fold higher luciferase activity than the promoterless pGL3-Basic expression vector in IEC-18 cells. However, a deletion from −302 to −157 and −157 to −106 resulted in 74 and 94% loss of luciferase activity, respectively, suggesting the presence of a positive and a basal regulatory element within −302/−106 region. These results suggest that at least two positive regulatory elements at −1063/−943 and −302/−157, two negative response regions at −2723/−1128 and −530/−443, and a basal transcription element at −157/−106.

This region from −443 to −19 contains five putative Sp1 binding sites (see Fig. 3A). To further explore the Sp1 functional role in *Mrp3* gene regulation, the five consensus Sp1 sites were individually or combined mutated by site-directed mutagenesis (Fig. 3A). Figure 3 shows the luciferase activity of each mutant construct relative to the activity of the pGL3-Basic vector. All the various mutants were compared with the same length of wild-type constructs. It was found that specific mutation of the Sp1 (3) (pMS-443-3, pMS-302-3, and pMS-157-3) or combined the Sp1 (3) mutant (pMS-157-3, 4, pMS-157-3, 4, 5, and pMS-157-3, 5) results in almost complete reduction (more than 90%) in luciferase activity. Similar result has been obtained in further deletion mutant (pWT-106). However, specific mutation of the Sp1 (4) (pMS-157-4) or combined the Sp1 (4) mutant (pMS-157-4, 5) reduced its activity by 40–50%. These results demonstrate that the region Sp1 (3) and Sp1 (4) are essential for basal promoter activity. It is interesting to note that none of the other mutated constructs altered promoter activity more than 2.5-fold. It is possible that there are redundant regulatory elements in the promoter and that mutation of one element is compensated for by another element. In this study, it was indicated that there were multiple Sp1 sites of the proximal 5′-flanking region involved in combinatorial control of *Mrp3* gene transcription.

To clarify whether Sp1 or Sp3 could bind the Sp1 (3) and Sp1 (4) consensus sequence, we used the wild-type DNA probe (WT), and the sequence containing Sp1 site-specific mutations (see Fig. 4A) as substrates in EMSAs with nuclear extracts from IEC-18 and SL2 cells, which did not express Sp1 or Sp1-related proteins. Promega Sp1 consensus oligo-DNA was used as a positive control (Fig. 4A). It was shown that with the WT probe there were three retarded bands (A, B, and C) as the same as that with Sp1 oligonucleotide (Fig. 4A). However, with mismatched probe (MS), the three retarded bands disappeared. No complex is formed in the absence of cell nuclear extract. Supershift assays provided evidence for Sp1 (A) and Sp3 or Sp3-related protein (B and C) binding within nt −123/−105 (Fig. 4B). These studies indicate that Sp1, Sp3, or Sp3-related transcription factor(s) may involve in the regulation of the rat *Mrp3* gene.

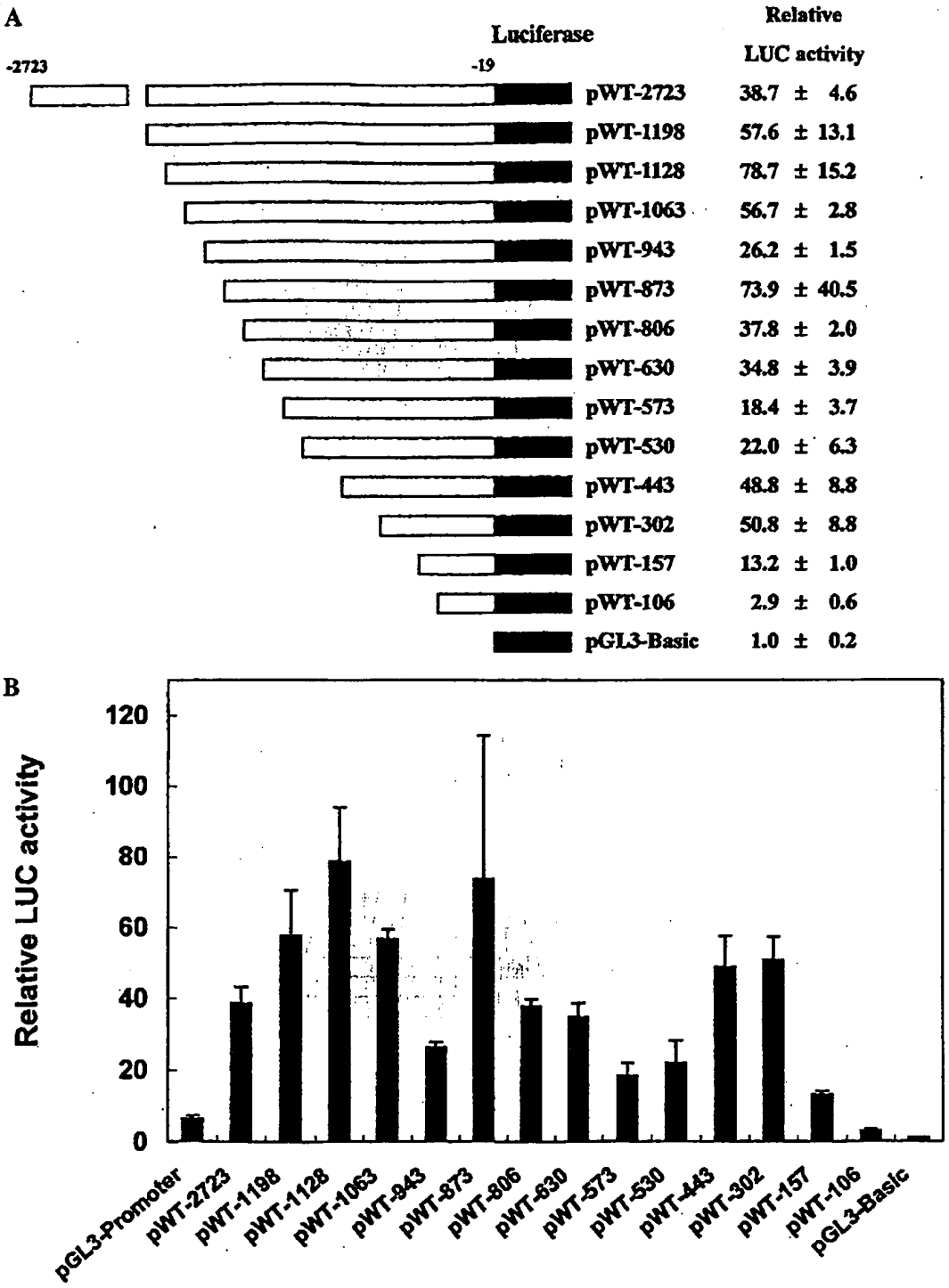


FIG. 2. 5'-Deletional analysis of the rat *Mrp3* gene in transient transfected IEC-18 cells. (A) Schematic representation of the structure of the 2.7-kb rat *Mrp3* gene promoter and its sequentially deleted fragments inserted upstream from coding region (indicated as a black box at the 3' end of each insert) in the pGL3-Basic vector (Promega). A series of 5' deletion constructs beginning at base −2723 with identical 3'-end (base −19) is shown. The numbers indicate the positions of the 5' ends of the deletion constructs relative to the first base in the translation initiation codon (ATG, +1). The name of the plasmid denotes the 5'-upstream extension of the insert DNA. (B). These deletion constructs were analyzed by transient transfections in IEC-18 cells. Cells were cotransfected with a total of 0.6 μg of the indicated *Mrp3*-luciferase reporter gene and pRL-TK, contains a *Renilla* luciferase, as an internal control. Luciferase activity was assessed 72 h after transfection. The luciferase activities for each reporter construct were normalized to the *Renilla* luciferase activities. Activities are reported as fold induction over that seen in the cells transfected using the pGL3-Basic vector. Results represent means ± SD for three or more transfections assayed in triplicate.

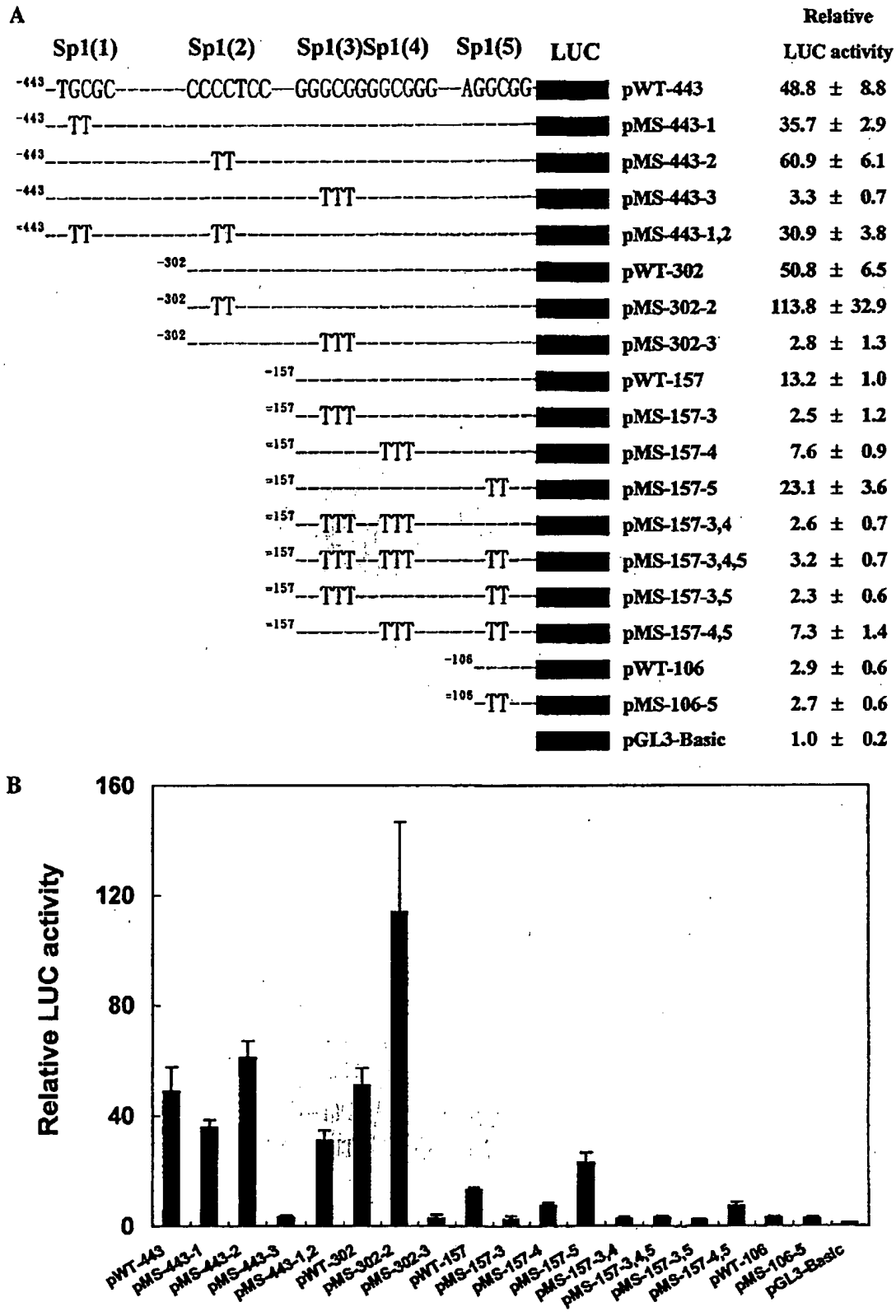


FIG. 3. Effect of site-specific mutations in Sp1 sites and Sp1 deletion mutants on luciferase activity. (A) Site-directed mutagenesis constructs of Sp1 consensus sequences in the 5′ flanking region from −443 to −19 bp of Mrp3 gene are shown. Five putative Sp1 recognition sites are shown. (B) Transcriptional analysis of Sp1 site-mutated and Sp1 deletion mutants. The luciferase activities of each deletion and mutation construct were normalized to the *Renilla* luciferase activities. Activities are expressed as fold induction over pGL3-Basic vector. Each value represents the mean ± SD of at least three independent sets of transfection experiments in triplicate.

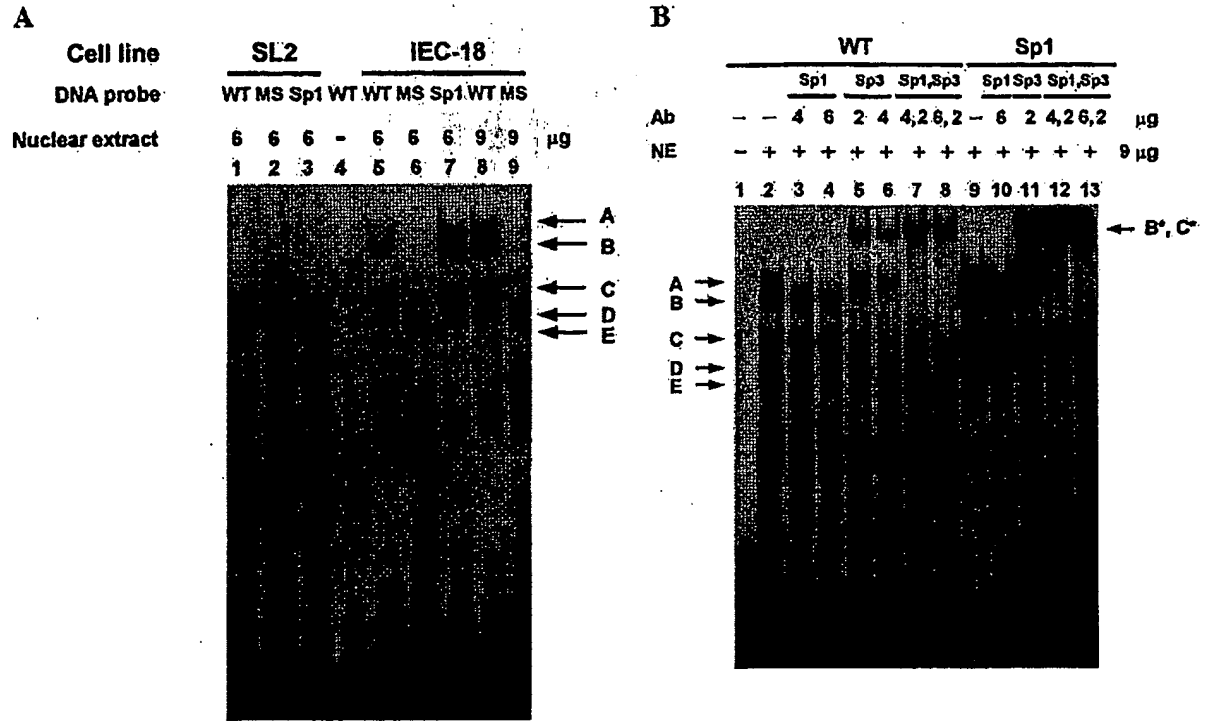


FIG. 4. Electrophoretic mobility shift assays the nuclear factors binding to the *Mrp3* gene promoter. (A) These DNA probes were labeled with [γ-^{32}P]ATP by the T4 polynucleotide kinase and incubated with nuclear extracts from IEC-18 cells. Protein–DNA complexes with wild-type probe (WT) are labeled A, B, C, D, and E. (B) In supershift assay Sp1 monoclonal antibodies or Sp3 polyclonal antibodies (Santa Cruz) were added to nuclear extracts prior to the addition of labeled WT probe. * represents a supershift band by Sp1 or Sp3 antibody.

Functional analyses of site-directed mutagenesis constructs of Sp1 sites were demonstrated that the region within −123/−105 was required for basal promoter activity. Moreover, supershift assays demonstrated that Sp1 and Sp3 bind to this region. The results indicate that Sp1 and Sp3 may involve in rat *Mrp3* gene regulation. These data were consistent with the other reports that the regulation of MRP family transcription required GC-rich sites located within the first 200 bases of its promoter and had similar DNA context structure (17–21).

To date, three other Sp1-related proteins, Sp2, Sp3, and Sp4 have been identified (22, 23). Sp3, like Sp1, is expressed in various cells. Sp3 can serve not only as a transcriptional activator but also as a repressor of Sp1-mediated transcription depending on cellular and promoter contexts (24, 25). The details of how these transcription factors are involved in the *Mrp3* gene activity of IEC-18 cells may help to explain the different expression of Mrp3 in the liver and intestine. Moreover, the interactions between Sp1, Sp3, and other transcription factors, like CBF/NF-Y (26), FTF (27), nuclear receptor (PXR, FXR, and CAR) (28), AP-1 (21), and p53 (29), may provide tissue specificity in MRP family gene systems and will be particularly interesting. Further studies are required to elucidate these ideas.

ACKNOWLEDGMENT

This work was supported by Grant NSC89-2320-B006-142 from the National Science Council of the Republic of China (Taipei).

REFERENCES

1. Kool, M., de Hass, M., Scheffer, G. L., Scheper, R. J., van Eijk, M. J., Juijn, J. A., Baas, F., and Borst, P. (1997) Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5, homologues of the multidrug resistance-associated protein gene (MRP1), in human cancer cell lines. *Cancer Res.* **57**, 3537–3547.
2. Hirohashi, T., Suzuki, H., and Sugiyama, Y. (1999) Characterization of the transport properties of cloned rat multidrug resistance-associated protein 3 (MRP3). *J. Biol. Chem.* **274**, 15181–15185.
3. Zeng, H., Bain, L. J., Belinsky, M. G., and Kruh, G. D. (1999) Expression of multidrug resistance protein-3 (multispecific organic anion transporter-D) in human embryonic kidney 293 cells confers resistance to anticancer agents. *Cancer Res.* **59**, 5964–5967.
4. Kool, M., van der Linder, M., de Haas, M., Scheffer, G. L., de Vree, J. M., Smith, A. J., Jansen, G., Peters, G. J., Ponne, N., Scheper, R. J., Elferink, R. P., Baas, F., and Borst, P. (1999)

MRP3, an organic anion transporter able to transport anticancer drugs. *Proc. Natl. Acad. Sci. USA* **96**, 6914–6919.

5. Hirohashi, T., Suzuki, H., Ito, K., Ogawa, K., Kume, K., Shimizu, T., and Sugiyama, Y. (1998) Hepatic expression of multidrug resistance-associated protein-like proteins maintained in Eisai hyperbilirubinemic rats. *Mol. Pharmacol.* **53**, 1068–1075.
6. Kiuchi, Y., Suzuki, H., Hirohashi, T., Tyson, C. A., and Sugiyama, Y. (1998) cDNA cloning and inducible expression of human multidrug resistance associated protein 3 (MRP3). *FEBS Lett.* **433**, 149–152.
7. Laitinen, M., and Watkins, J. B. (1986) Mucosal biotransformation. *In* Gastrointestinal Toxicology (Rozman, K., and Hanninen, O., Eds.), pp. 169–192, Elsevier, Amsterdam.
8. Lin, J. H., Chiba, M., and Baillie, T. A. (1999) Is the role of the small intestine in first-pass metabolism overemphasized? *Pharmacol. Rev.* **51**, 135–157.
9. Peters, W. H. M., Nagengast, F. M., and van Tongeren, J. H. M. (1989) Glutathione *S*-transferase, cytochrome P450, and uridine 5'-diphosphate-glucuronosyltransferase in human small intestine and liver. *Gastroenterology* **96**, 783–789.
10. Radominska-Pandya, A., Littl, J. M., Pandya, J. T., Tephly, T. R., King, C. D., Barone, G. W., and Raufman, J. P. (1998) UDP-glucuronosyltransferases in human intestinal mucosa. *Biochim. Biophys. Acta* **1394**, 199–208.
11. Koster, A. S., and Noordhoek, J. (1983) Glucuronidation in isolated perfused rat intestinal segments after mucosal and serosal administration of 1-naphthol. *J. Pharmacol. Exp. Ther.* **226**, 533–538.
12. Wollenberg, P., and Rummel, W. (1985) Influence of phosphate, sulfonic, and sulfamic acids on sulfoconjugate release in the vascularly perfused mouse small intestine. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **329**, 195–200.
13. Donner, M. G., and Keppler, D. (2001) Up-regulation of basolateral multidrug resistance protein 3 (Mrp3) in cholestatic rat liver. *Hepatology* **34**, 351–359.
14. Hirohashi, T., Suzuki, H., Takikawa, H., and Sugiyama, Y. (2000) ATP-dependent transport of bile salts by rat multidrug resistance-associated protein 3 (Mrp3). *J. Biol. Chem.* **275**, 2905–2910.
15. Chen, B. K., and Chang, W. C. (2000) Functional interaction between c-Jun and promoter factor Sp1 in epidermal growth factor-induced gene expression of human 12(*S*)-lipoxygenase. *Proc. Natl. Acad. Sci. USA* **87**, 7477–7481.
16. Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolodny, N., L., and Kolchanov, N. A. (1998) Databases on transcriptional regulation: TRANSFAC, TRRD, and COMPEL. *Nucleic Acids Res.* **26**, 264–370.
17. Kauffmann, H. M., and Schrenk, D. (1998) Sequence analysis and functional characterization of the 5'-flanking region of the rat multidrug resistance protein 2 (mrp2) gene. *Biochem. Biophys. Res. Commun.* **245**, 325–331.
18. Tanaka, T., Uchiumi, T., Hinoshita, E., Inokuchi, A., Toh, S., Wada, M., Takano, H., Kohno, K., and Kuwano, M. (1999) The human multidrug resistance protein 2 gene: Functional characterization of the 5'-flanking region and expression in hepatic cells. *Hepatology* **30**, 1507–1512.
19. Fromm, M. F., Leake, B., Roden, D. M., Wilkinson, G. R., and Kim, R. B. (1999) Human MRP3 transporter: Identification of the 5'-flanking region, genomic organization and alternative splice variants. *Biochim. Biophys. Acta* **1415**, 369–374.
20. Takada, T., Suzuki, H., and Sugiyama, Y. (2000) Characterization of 5'-flanking region of human MRP3. *Biochem. Biophys. Res. Commun.* **270**, 728–732.
21. Kurz, E. U., Cole, S. P., and Deeley, R. G. (2001) Identification of DNA–protein interactions in the 5' flanking and 5' untranslated regions of the human multidrug resistance protein (MRP1) gene: Evaluation of a putative antioxidant response element/AP-1 binding site. *Biochem. Biophys. Res. Commun.* **285**, 981–990.
22. Hagen, G., Müller, S., Beato, M., and Suske, G. (1992) Cloning by recognition site screening of two novel GT box binding proteins: A family of Sp1 related genes. *Nucleic Acids Res.* **20**, 5519–5525.
23. Kingsley, C., and Winoto, A. (1992) Cloning of GT box-binding proteins: A novel Sp1 multigene family regulating T-cell receptor gene expression. *Mol. Cell. Biol.* **12**, 4251–4261.
24. Hagen, G., Müller, S., Beato, M., and Suske, G. (1994) Sp1-mediated transcriptional activation is repressed by Sp3. *EMBO J.* **13**, 3843–3851.
25. Suske, G. (1999) The Sp-family of transcription factors. *Gene* **238**, 291–300.
26. Kauffmann, H. M., Vorderstemann, B., and Schrenk, D. (2001) Basal expression of the rat, but not of the human, multidrug resistance protein 2 (MRP2) gene is mediated by CBF/NF-Y and Sp1 promoter-binding sites. *Toxicology* **167**, 25–35.
27. Inokuchi, A., Hinoshita, E., Iwamoto, Y., Kohno, K., Kuwano, M., and Uchiumi, T. (2001) Enhanced expression of the human multidrug resistance protein 3 by bile salt in human enterocytes. *J. Biol. Chem.* **276**, 46822–46829.
28. Kast, H. R., Goodwin, B., Tarr, P. T., Jones, S. A., Anisfeld, A. M., Stoltz, C. M., Tontonoz, P., Kliewer, S., Willson, T. M., and Edwards, P. A. (2001) Regulation of multidrug resistance-associated protein 2 (MRP2;ABCC2) by the nuclear receptors PXR, FXR, and CAR. *J. Biol. Chem.*, in press.
29. Sullivan, G. F., Yang, J. M., Vassil, A., Yang, J., Bash-Babula, J., and Hait, W. N. (2000) Regulation of expression of the multidrug resistance protein MRP1 by p53 in human prostate cancer cells. *J. Clin. Invest.* **105**, 1261–1267.

ELSEVIER

Gene 282 (2002) 75–86

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Characterization of the promoter of human extracellular matrix metalloproteinase inducer (EMMPRIN)

Liang Liang, Terry Major, Thomas Bocan*

Department of Cardiovascular Pharmacology, Pfizer Global Research and Development, Ann Arbor Laboratories, Pfizer Inc., 2800 Plymouth Road, Ann Arbor, MI 48105, USA

Received 13 July 2001; received in revised form 21 October 2001; accepted 20 November 2001
Received by R. Di Lauro

Abstract

Monocyte-derived macrophages play a central role in atherosclerotic lesion formation and potentially plaque destabilization by expression of matrix metalloproteinases (MMPs); however, mechanisms associated with stimulating MMP production are not clearly understood. EMMPRIN, which is expressed by human cancer cells and macrophages, present in human, mouse and rabbit atherosclerosis and noted to induce MMPs may be involved. A DNA fragment containing 1797 bp 5′ upstream of the EMMPRIN gene and the transcription start site was generated by polymerase chain reaction and cloned into a luciferase reporter gene vector, pGL3-basic. The relative luciferase activities driven by this 5′-upstream fragment and a series of deletion mutants were measured in transiently transfected human and mouse macrophage THP-1 and Raw264.7 cells, respectively. A fragment 471 bp upstream of the EMMPRIN coding region was sufficient to promote transcription, while a region from −1413 to −1024 bp suppressed activity. Further deletion analysis of the 471 bp fragment indicated that a 30 bp element from −142 to −112 bp, which contains binding sites for Sp1, AP1TFII and EGR-2, was important for EMMPRIN transcription in both THP-1 and Raw264.7 macrophages. Using electrophoretic mobility shift assays, the Sp1 element within 30 bp region specifically bound Sp1 and Sp3 transcription factors. Mutation of the Sp1 element at −122 to −116 bp of the EMMPRIN promoter significantly diminished promoter activity and formation of DNA-nuclear protein complex. Transient expression of Sp1 and/or Sp3 transcription factors in insect cells lacking the Sp family of transcription factors, stimulated EMMPRIN promoter activity in a synergistic manner. Together, these results indicate that both Sp1 and Sp3 associate with the functional Sp1 element on the EMMPRIN promoter and cooperate in the regulation of EMMPRIN gene expression in macrophages. © 2002 Published by Elsevier Science B.V.

Keywords: Sp1; Sp3; Macrophage; Matrix metalloprotein; Transcriptional regulation

1. Introduction

Monocyte-derived macrophages play a central role in the atherosclerotic lesion initiation, progression and regression (Libby and Clinton, 1993). Adhesion and invasion of monocytes into the intima are early events in atherosclerosis formation (Joris et al., 1983). Differentiation of monocytes to macrophages, formation of lipid-loaded macrophage foam cells with subsequent accumulation of foam cells within the arterial wall set the stage for progression of atherosclerosis and eventually formation of complicated plaques that may be prone to rupture (Ross, 1999). Plaque rupture, a cause of stroke and myocardial infarction, often occurs at the vulnerable shoulder regions containing macrophages (Lendon et al., 1991). Circumstantial evidence suggests that macrophage secretion and activation of matrix matalloproteinases (MMPs) may be involved in the plaque destabilization (Galis et al., 1994; Galis et al., 1995).

MMPs are a family of Zn^{2+}-containing enzymes that cleave most of the components of extracellular matrix (ECM). MMPs are associated with macrophages of atherosclerotic lesions, and are particularly located in the macrophage-rich shoulder area (Galis et al., 1994; Galis et al., 1995). Monocyte/macrophage MMPs, by cleaving many of the ECM structural proteins, have been reported to play a role in the migration of monocytes through the intima under pathological settings (Amorino and Hoover, 1998). In addition, macrophage-derived MMPs in the vulnerable shoulder area of atherosclerotic lesions may, at least

Abbreviations: EMMPRIN, extracellular matrix metalloproteinase inducer; MMP, matrix metalloproteinase(s); MT-MMP, membrane type MMP; ECM, extracellular matrix; EMSA, electrophoretic mobility shift assay; PCR, polymerase chain reaction; HMDM, human monocyte-derived macrophage

* Corresponding author. Tel.: +1-734-622-7383; fax: +1-734-622-4398.
E-mail address: thomas.bocan@pfizer.com (T. Bocan).

0378-1119/02/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S0378-1119(01)00847-2

partially, be responsible for degradation of ECM associated with plaque rupture (Galis et al., 1994). Therefore, regulation of MMP expression in monocyte/macrophages is a potentially important phenomenon in the atherosclerotic plaque formation and rupture.

Extracellular matrix metalloproteinase inducer (EMMPRIN) is one of the factors involved in the production and activation of MMPs (Guo et al., 1997; Sun and Hemler, 2001). EMMPRIN (also known as CD147, basigin, and tumor cell-derived collagenase stimulatory factor), is a member of the immunoglobulin superfamily (Biswas et al., 1995). EMMPRIN is a highly glycosylated transmembrane protein initially identified on the surface of human cancer cells and shown to stimulate adjacent stromal cells to produce several MMPs (Guo et al., 1998), including MMP-1, MMP-2, MMP-3, membrane type 1 MMP (MT1-MMP), and MT2-MMP (Guo et al., 1997; Kataoka et al., 1993; Prescott et al., 1989; Sameshima et al., 2000). We have previously found that EMMPRIN was expressed in human monocyte-derived macrophages and was upregulated upon monocyte macrophage differentiation. Macrophage EMMPRIN expression was correlated with MMP-9 upregulation and found to colocalize with macrophages in atherosclerotic lesions (Major et al., 2000). These data raise the possibility that monocyte/macrophage-expressed EMMPRIN, acting as in cancer cells, may play a central role in atherosclerotic lesion development, macrophage accumulation and MMP production.

Despite its potential role in atherosclerosis and cancer, the molecular regulation of EMMPRIN expression remains unclear. In the present study, we have characterized the transcriptional regulation of the EMMPRIN gene in macrophages and have found that the Sp1 and Sp3 transcription factors, by binding with Sp1 consensus element located at −122 to −116 bp of the EMMPRIN promoter, regulate EMMPRIN gene expression synergistically. These studies have for the first time characterized EMMPRIN gene regulation in the macrophage and may serve as the basis for studying the regulation of arterial wall MMP production.

2. Materials and methods

2.1. Materials

PMA and chemicals were purchased from Sigma (St Louis, MO) unless otherwise noted.

2.2. Cell lines and cell culture

The THP-1, Raw264.7, Schneider's *Drosophila melanogaster* cell line 2 (SL-2) and 293 cells were purchased from ATCC (Rockville, MD). THP-1 cells were grown in RPMI media 1640 supplemented with 25 mM HEPES, Glutamax (Gibco BRL, Rockville, MD), 100 units/ml penicillin, 100 μg/ml streptomycin (Gibco BRL), and 10% fetal bovine serum (FBS) (HyClone Laboratories, Logan, UT). Raw246.7, Human breast carcinoma cells (MCF-7) and CHO cells were grown in Dulbecco's modified Eagle medium supplemented with 4500 mg/l D-glucose, L-glutamine, pyridoxine hydrochloride (Gibco BRL), 100 units/ml penicillin, 100 μg/ml streptomycin (Gibco BRL), and 10% fetal bovine serum (FBS) (HyClone Laboratories, Logan, UT). 293-cells were maintained by following the ATCC growth protocol. Cells were incubated in the humidified incubator equilibrated with 5% CO_2, at 37 °C. Suspension cell cultures of THP-1 cells were maintained at a density less than 5×10^5 cells/ml and cell differentiation was achieved by treating the THP-1 cells with 40 nM PMA (Sigma) for 48 h. SL-2 was maintained at 24 °C in Schneider's Insect Medium (Gibco BRL) supplemented with 10% heat-inactivated fetal bovine serum (HyClone Laboratories) and 100 units/ml penicillin, 100 μg/ml streptomycin (Gibco BRL). Human monocytes, supplied by Advanced Biotechnology (Columbia, MD), were maintained in DMEM containing 10% FBS, 1% penicillin/streptomycin. Human monocytes were differentiated into macrophages (HMDM) by culturing in DMEM containing 10% FBS, 1% penicillin/streptomycin and 1 ng/ml GM-CSF (R&D Systems, Minneapolis, MN) for 10 days. On day 10, the cells were washed with phosphate-buffered saline (PBS) and the experiments were conducted.

2.3. RNA isolation and Northern analysis

Total RNA was isolated by the guanidium isothiocyanate method (Chomczynski and Sacchi, 1987). Fifteen micrograms of total RNA were electrophoresed in a 1% formaldehyde/agarose gel and blotted onto a nylon membrane (Ambion, Austin, TX) in 3 M NaCl, 0.3 M sodium citrate (20× SSC) (Gibco BRL) by capillary transfer. The Northern blots were UV cross-linked by UV Stratalinker1800 (Stratagene, La Jolla, CA) and prehybridized using ExpressHyb Hybridization Solution (Clontech, Palo Alto, CA) at 68 °C for 30 min in a Hybridizer 700 (Stratagene). The blots were hybridized at 68C overnight with fresh ExpressHyb Solution containing radiolabeled [α-^{32}P]dCTP cDNA probes. Blots were washed in 2× SSC, 0.05% SDS solution for 30 min at room temperature and then washed in 0.1× SSC, 0.1% SDS solution with continuous shaking at 50 °C for 40 min. The blots were directly exposed on a Storm 860 PhosphorImager and the results were analyzed by ImageQuant software (Molecular Dynamics). The cDNA probe for human and mouse EMMPRIN (591 bp fragment (+119 to 709 bp) and 532 bp fragment (+672 to 1203 bp), respectively) were generated by PCR and subcloned into pCRII vector for synthesis and fragmented by *Eco*RI digestion. The 0.9 kb cDNA for human S9 ribosomal was purchased from Clontech. The cDNAs were labeled by [α-^{32}P]dCTP using Rediprime II random prime labeling system (Amersham Pharmacia Biotech, Buckinghamshire, UK) following the manufacturer's protocol.

2.4. Human EMMPRIN promoter cloning and DNA sequencing

An 1812 bp DNA fragment (−1797 to +15 bp) upstream from the translation initiation site of EMMPRIN was PCR amplified with Advantage −GC Genomic PCR kit (Clontech) which contains the polymerase mix with 3′–5′ proofreading activity. Human DNA libraries from a Genomewalker kit (Clontech) was used as the template. The forward primer (5′-GCTCTAGAGGATCCCACCCTTCTTTA) and the reverse primer with an *Nhe*I site (5-GCTAGCAGCGCAGCCGCCATGATTC) used in the PCR reaction were designed based on information from the human genomic data base describing the sequence of human chromosome 19. The fragment was gel-purified and subcloned into pCR-Blunt II–TOPO vector (Invitrogen, Carlsbad, CA). The *Nhe*I and *Kpn*I restriction enzyme digested fragment which contains the EMMPRIN promoter was subcloned into a Luciferase reporter gene vector pGL3-basic (Promega, Madison, WI) and the DNA was sequenced with an automatic DNA sequencer (MWG Biotech, High Point, NC). The potential transcription factor binding sites were searched using TESS computer program.

2.5. Mutagenesis of the human EMMPRIN promoter

A luciferase reporter plasmid driven by 1797 bp of the human EMMPRIN promoter was used in this study and was referred to as −1797 bp. 1797 bp EMMPRIN promoter plus +15 bp of EMMPRIN gene/pCR-Blunt II–TOPO was used as a template to synthesize the serial deletion mutants −58 bp, −471 bp, −1024 bp, and −1413 bp. These mutants were generated by *NotI*, *SacI*, *BamHI*, *PstI* enzyme digestion, respectively, and self-ligated. The plasmids were transformed into TOP 10 competent cells (Invitrogen), extracted, *XhoI/HindIII* digested for −58 bp, *KpnI/NheI* digested for −471 bp and −1024 bp, *NheI/HindIII* digested for −1413 bp, gel-purified, and subcloned into the promoterless luciferase reporter gene pGL3-basic (Promega) with *XhoI/HindIII*, *KpnI/NheI*, Xba/*HindIII* enzyme sites. Further deletion mutagenesis of the proximal sequence (−471 to −58 bp) of the human EMMPRIN promoter was generated by PCR. Forward primers with *Kpn*I site −112 bp, 5′-GGTACCAGATGACGCCGTGCGTGC, −142 bp, 5′-GGTACCCC-GGCGTCCCCGGCGCTC, −197 bp, 5′-GGTACCCTG-CCGGAGCCGGCGCGT, −267 bp, 5′-GGTACCGCAG-GCCCCACTCCCGTT, and the reverse primer with *Nhe*I site 5′-GCTAGCAGCGCAGCCGCCATGATTC were used to perform the PCR reaction with −471 bp construct as a template (Advantage −GC Genomic PCR kit, Clontech). PCR products were gel-purified, and subcloned into pCR-Blunt II–TOPO vector (Invitrogen, Carlsbad, CA). The constructs with pCR-Blunt II–TOPO vector were digested with enzyme *KpnI/NheI*, the DNA fragments of EMMPRIN promoter were gel-purified and subcloned into pGL3-basic vector with *KpnI/NheI* site. Site-directed mutagenesis was utilized to mutate the potential SP1 transcription element in the −122 to −116 bp of the human EMMPRIN promoter. The forward primer with *Kpn*I site for PCR, which introduced a mutation (underline), was 5′-GGTACCGGCGTCCCCGGCGCTCGAGATCTACCC-GAGATGACG and the reverse primer with *Nhe*I site was 5′-GCTAGCAGCGCAGCCGCCATGATTC. PCR reaction was performed with −142 bp as template. The PCR product was subcloned into pCR-Blunt II–TOPO vector and then subcloned into pGL3-basic vector with *KpnI/NheI* sites as described above.

2.6. Transient transfection and dual luciferase assay

Transfection of the THP-1 cells was performed by electroporation, as previously described (Liao et al., 1997). Briefly, 5×10^6 THP-1 cells in 400 μl of RPMI-1640 medium were transfected with 18 μg of the EMMPRIN promoter/pGL3-basic constructs (deletion and site-directed mutation constructs) and 2 μg of pRL-tk gene (Promega) by electroporation with a Bio-Rad Gene Pulser set at 250 V, 960 μF. One hour after transfection, THP-1 cells were treated with 40 nM PMA for 40 h and then harvested for the luciferase assay. Transfection of the Raw264.7 cells was performed by lipofectamine (LF) 2000 (Gibco BRL) following the manufacturer's protocol. Briefly, 2×10^5 cells per well were plated in a 24-well plate the day before transfection. For each well of cells to be transfected, 5 μg of the EMMPRIN promoter/pGL3-basic constructs (deletion and site-directed mutation constructs) and 0.05 μg of pRL-tk gene were mixed with LF2000 reagent in OPTI-MEM I medium (Gibco BRL) and incubated for 20 min at room temperature. The DNA-LF2000 reagent complexes were added to the cells and 1 ml of growth media was added to each well of cells after 6 h of transfection. The cells were harvested for the luciferase assay after 24 h incubation. Transfections were performed in triplicate for each construct.

Dual luciferase assays were designed for measuring fire fly luciferase activity driven by the EMMPRIN promoter and renilla luciferase activity driven by co-transfected pRL-tk, which served as an internal control and the baseline response, within a single sample. The activity of EMMPRIN promoter (firefly luciferase activity) was normalized by the activity of Renilla luciferase to minimize the experimental variability caused by differences in cell viability or transfection efficiency. Cells were washed with PBS and lysed with 100 μl of 1× Passive Lysis Buffer (Promega) for 15 min. Twenty microliters of cell lysate were transferred to each well of 96-well plate and both fire fly and Renilla luciferase activities were measured in triplicate for all samples with the promega dual luciferase assay system and a microplate Luminometer LB96V (EG&G, Berthold). The luminometer was programmed to perform a 2-sec pre-measurement delay, followed by a 10-sec measurement period for each reporter assay. One hundred microliters of of Luciferase

Assay Reagent II (Promega) was automatically injected and firefly luciferase activities were measured. Then, 100 μl of Stop & Glo Reagent (Promega) was injected and the Renilla luciferase activities were read. The firefly luciferase activity of each sample was normalized to renilla luciferase activity and each sample was then compared to the maximal activity.

For transfection of SL2 cells, one day prior to transfection, cells were plated into 6-well plates at a density of 2×10^6 cells/well. Cells were transfected by lipofectamine (LF) 2000 (Gibco BRL) following the manufacturer's protocol. Each well received 2.5 μg of the −142 bp EMMPRIN promoter construct and variable amounts of the expression plasmids pPacSp1 and pBKCMVSp3. Salmon Sperm DNA was used to adjust the variable amount of expression plasmid. The medium was changed 24 h after addition of the DNA and then the cells were harvested 48 h after transfection. Cells were lysed with 200 μl/well of passive lysis buffer (Promega) for luciferase assay. Luciferase activities were normalized by total protein concentration as determined by the BioRad protein assay.

2.7. Nuclear extracts and electrophoretic mobility shift assay (EMSA)

Nuclear extracts were prepared from THP-1 cells treated with 40 nM PMA for 48 h and Raw264.7 cells. Cells were washed twice with ice-cold PBS, once with cold PBS plus 1 mM sodium orthovanadate (Na_3VO_4) plus 5 mM NaF, once with Hypotonic Buffer (20 mM HEPES, 20 mM NaF, 1 mM Na_3VO_4, 1 mM phosphate buffer (pH 7.4), 1 mM EGTA (pH 8.5), 1 mM DTT, 0.5 mM PMSF, 0.125 μM okadaic acid, one Complete Pellet from Boehringer Mannheim). Hypotonic buffer containing 0.2% NP-40 was added to lyse the cell pellet. After 5 min incubation on ice, nuclei were pelleted by centrifugation at 14,000 rpm for 1 min at 4 °C. Nuclei were resuspended in High Salt Buffer (Hypotonic Buffer contains 0.42 M NaCl, 20% glycerol) and rocked at 4 °C for 30 min. The nuclear extracts (supernatant) were collected after centrifugation at 4 °C at 14,000 rpm for 20 min, aliquoted and stored at −70 °C. The protein concentration of an aliquot was determined using the Bio-Rad protein assay reagent.

The sequences of the sense strand of double-stranded oligonucleotides used as probes or competitors in EMSA with the mutation underlined were as follows: −142 to −112 wild-type, 5′-CCGGCGTCCCCGGCGCTCGCCCCGCCCCCG, −142 to −112 mutant, 5′-CCGGCGTCCCCGGCGCTCGAGATCTACCCG, Sp1 consensus, 5′-ATTCGATCGGGGCGGGGCGAGC, Sp1 mutant, 5′-ATTCGATCGGTTCGGGGCGAGC, AP1 consensus, 5′-CGCTTGATGAGTCAGCCGGAA, TFII consensus, 5′-GCAGAGCATATAAGGTGAGGTAGGA, and EGR consensus, 5′-GGATCCAGCGGGGGCGAGCGGGGGCGA. The probes were labeled with T4 Polynucleotide Kinase (Gibco BRL) and [γ-^{32}P]ATP. The DNA binding reaction was carried out in a 20-μl reaction mixture containing Gel Shift Binding Buffer (Promega), 10 μg of nuclear extract, and 20,000 cpm of indicated ^{32}P-labeled probe with or without 100-fold amount of unlabeled competitor oligonucleotide. The reaction mixture was incubated at room temperature for 20 min and loaded onto a 5% polyacrylamide gel. For the supershift assay, the DNA binding reaction mixture was incubated with 1μl of 200 μg/0.1 ml rabbit polyclonal antibodies against Sp1, Sp2, Sp3, Sp4 (Santa Cruz Biotechnology, Santa Cruz, CA) for 30 min at room temperature and load on a 4% gel. The bound and free DNA were resolved by electrophoresis at 300 V in 0.5× TBE (50 mM Tris–HCl, 2 mM EDTA). Dried gels were exposed to both Kodak XAR film and Storm 860 PhosphorImager.

2.8. Immunoblot analysis of nuclear extracts

Twenty micrograms of nuclear extracts from human monocytes or human monocyte-derived macrophages were loaded on 4–12% NuPAGE gel (Invitrogen). After transfer to the membrane, the nuclear extracts were immunoblotted with 1 μg/ml rabbit polyclonal antibody against Sp1 (Santa Cruz Biotechnology). Detection of the immune complexes was carried out with the Western blot detection system–Super Signal Chemiluminescent Substrate for Western Blotting (Pierce, Rockford, IL). The membrane was stripped in 2 M of glycine (pH 2.5) for 30 min at room temperature and reprobed with 1 μg/ml rabbit polyclonal antibody recognizing Sp3 (Santa Cruz Biotechnology).

2.9. Statistical analysis

Levels of significance for comparisons between samples were determined using one-tailed Student's *t*-test.

3. Results

3.1. EMMPRIN expression in human and mouse macrophages

As illustrated in Fig. 1, mRNA of human and mouse EMMPRIN was expressed in THP-1 and Raw264.7 macrophages. Northern blot analysis using 591 bp probe against human EMMPRIN showed that human EMMPRIN mRNA was detected in human PMA-differentiated THP-1 macrophages (Fig. 1A, lane 3), human kidney 293 cells (Fig. 1A, lane 2), and positive control human MCF-7 cancer cells (Fig. 1A, lane 4). However, human EMMPRIN DNA probe was not able to detected hamster EMMPRIN in CHO cells. Northern blot analysis using a 532 bp probe against mouse EMMPRIN detected EMMPRIN mRNA in Raw264.7cells (Fig. 1B, lane 2) at the expected size. Mouse kidney, which is known to express EMMPRIN, was used as a positive control (Fig. 1B, lane 3). S9 ribosomal message was measured as an internal control in these experiments. Expression of EMMPRIN mRNA in THP-1 and Raw264.7

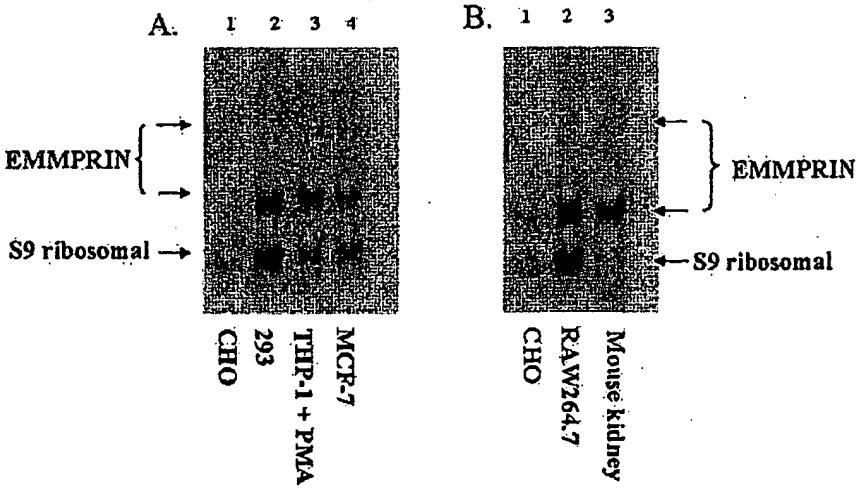


Fig. 1. EMMPRIN mRNA's expression in human (A) and mouse (B) macrophages. Fifteen micrograms of of total RNA isolated from CHO, 293, THP-1 with treatment of 40 nM PMA, MCF-7 cells, Raw264.7 cells and mouse kidney tissue were loaded in each lane of northern analysis. The blot was hybridized with the ^{32}P-labeled human (A) and mouse (B) EMMPRIN cDNA fragments and rehybridized with the human S9 ribosomal cDNA, a qualitative indicator of loading. The mRNA for EMMPRIN and S9 ribosomal were indicated by arrow.

cells not only confirmed that the expression machinery existed in these macrophage cell lines but also allowed us to assess the molecular basis of transcription in these cells.

3.2. Sequence of the human EMMPRIN promoter

The reported 942 bp 5′ flanking sequence and the transcription start site of EMMPRIN gene (Guo et al., 1998) were used as the basis for the more extensive cloning of the promoter. An 1812 bp DNA fragment containing −1797 bp upstream of initial codon of human EMMPRIN gene was cloned into pCR-Blunt II–TOPO vector by a PCR method and subsequently subcloned into a promoterless luciferase reporter vector pGL3-basic. The sequence of the DNA fragment is shown in Fig. 2 and the transcription initiation site, TATA box and some of the potential transcription factor binding elements on the proximal region are identified.

3.3. Mapping of the EMMPRIN promoter.

To identify sequences important for promoter activity, a series of 5′ deletion constructs were made by restriction enzyme digestion. 1812 bp (−1797 to +15 bp) DNA fragment containing the human EMMPRIN promoter in pCR-Blunt II–TOPO vector was used as wild-type to generate −1413 bp, −1024 bp, −471 bp, −58 bp of human EMMPRIN promoter deletion constructs by enzyme digestion. Both wild-type and deletion constructs were subsequently subcloned into the promoterless fire fly luciferase reporter vector pGL3-basic. These constructs are diagrammatically represented in Fig. 3A. The luciferase activities of these EMMPRIN promoter constructs were measured and normalized by renilla luciferase activities (internal control). The relevant luciferase activity was plotted as a percentage of the maximal activity. As shown in Fig. 3B,C, the luciferase activity of the −1413 to +15 bp region within the EMMPRIN promoter was decreased 5-fold in THP-1 cells and nearly ablated in Raw264.7 cells when compared to that of the −1024 to +15 bp EMMPRIN promoter region. These results suggest that there is a negative regulatory element within the −1413 to −1024 bp region. Constructs containing the −471 bp to +15 bp region of the EMMPRIN promoter drove maximal luciferase activity, while further deletion of the −471 to −58 bp regions in THP-1 or Raw264.7 cells caused a 10-fold reduction or complete loss of the luciferase activities, respectively, suggesting that elements promoting EMMPRIN gene transcription are present in this region. The constructs within the −58 to +15 bp region of the EMMPRIN promoter had essentially no luciferase activities.

To further define the transcriptional elements promoting EMMPRIN gene expression, we made a series of deletions by PCR to truncate different potential transcription factor binding sites within the −471 to −58 bp region. These constructs are diagrammatically shown in Fig. 4A. Each construct and pRL-tk were cotransfected into THP-1 and Raw264.7 cells as described above, and then the relative luciferase activity was determined (see Fig. 4B,C). Deletion of the −471 to −142 bp region of EMMPRIN promoter had no effect on luciferase activity in either cells. Further deletion of −142 to −112 bp of EMMPRIN promoter significantly decreased luciferase activity, suggesting that this 30 bp was required for transcription. The potential transcription factor binding elements in this region were searched using a TESS computer program and found that Sp1, AP1, EGR2, and TF II binding sites overlapped within this 30 bp GC-rich region (Fig. 2).

Fig. 2. The sequence of 1797 bp 5′-flanking region of EMMPRIN gene. +1 indicates the translation initiation site. The transcription start site and TATA box are indicated by arrow and box, respectively. Some consensus sequences for binding with transcription factor in the proximal region are underlined.

3.4. Binding of the EMMPRIN promoter

To identify the transcriptional elements that form a complex with specific transcription factors, nuclear extracts from PMA-treated THP-1 macrophage or Raw264.7 cells were analyzed by EMSA using ^{32}P labeled oligonucleotides spanning the 30 bp element from −142 to −112 bp. Two major complexes were detected in both THP-1 and Raw264.7 macrophages (Fig. 5A, lanes 2 and 7). To determine the transcription factor binding elements, excess of unlabeled Sp1, AP1, TF-II, and EGR2 consensus oligonucleotides were used as competitors to compete away the ^{32}P-labeled DNA–protein complexes (Fig. 5A, lanes 3–6, 8–11). Both DNA–protein complexes were competed away by Sp1 consensus oligonucleotides in both cells (Fig. 5A, lanes 3 and 8), indicating that a Sp1 element in this 30 bp region was important for binding with the transcription factors in the nuclear extracts. Nuclear extracts from HeLa cells were used as controls (Fig. 5A, lane 12) while labeled probe without nuclear extracts was used as a negative control (Fig. 5A, lane 1). To confirm this finding and test the specificity of this binding, a mutated Sp1 oligonucleotide (mSp1,

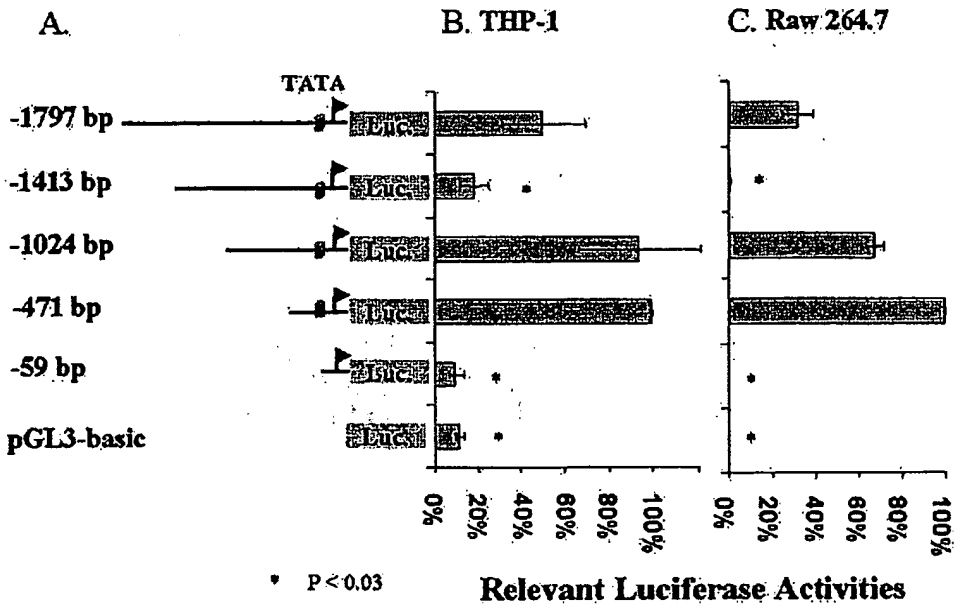


Fig. 3. (A) Diagram of the various EMMPRIN promoter constructs. The EMMPRIN promoter region with transcription start site (►) and TATA box and the extent of the 5′-truncations of EMMPRIN promoter are shown. The promoter fragments were ligated to the luciferase reporter gene in the forward orientation. (B,C) Activity of EMMPRIN promoter in human (B) and mouse (C) macrophages. 5′-deletion constructs of human EMMPRIN promoter were transiently cotransfected into THP-1 cells (B) and Raw264.7 cells (C) with pRL-tk gene, an internal control. The luciferase activity of each sample was normalized by the expression of the internal control and percentage of maximal activity was obtained by comparing the activity of each sample to the activity of −471 bp EMMPRIN promoter, which was 100% of the activity. Data are expressed as the mean ± SEM. Statistical analysis was performed comparing the activity of each construct to that of −471 bp construct that achieved the maximal activity. *$P < 0.03$.

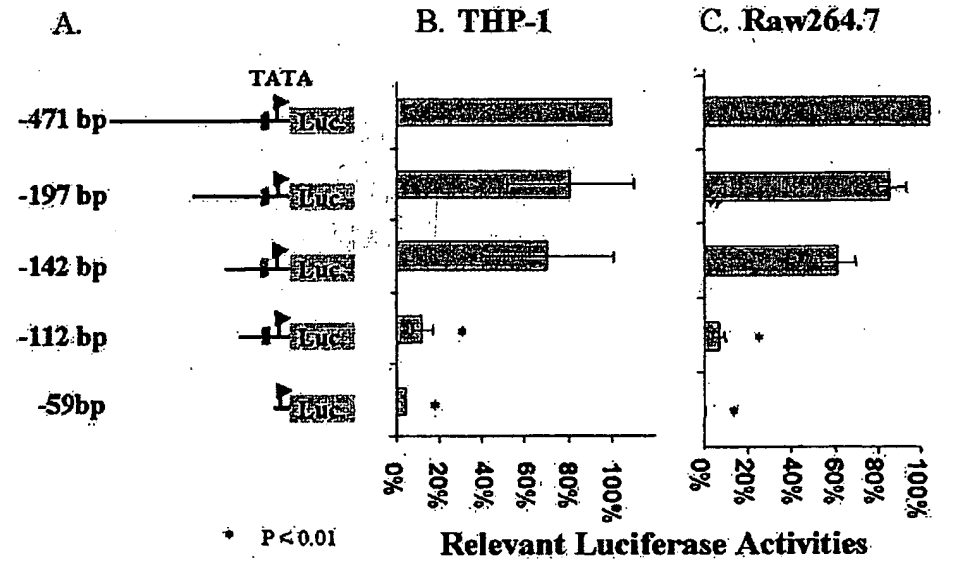
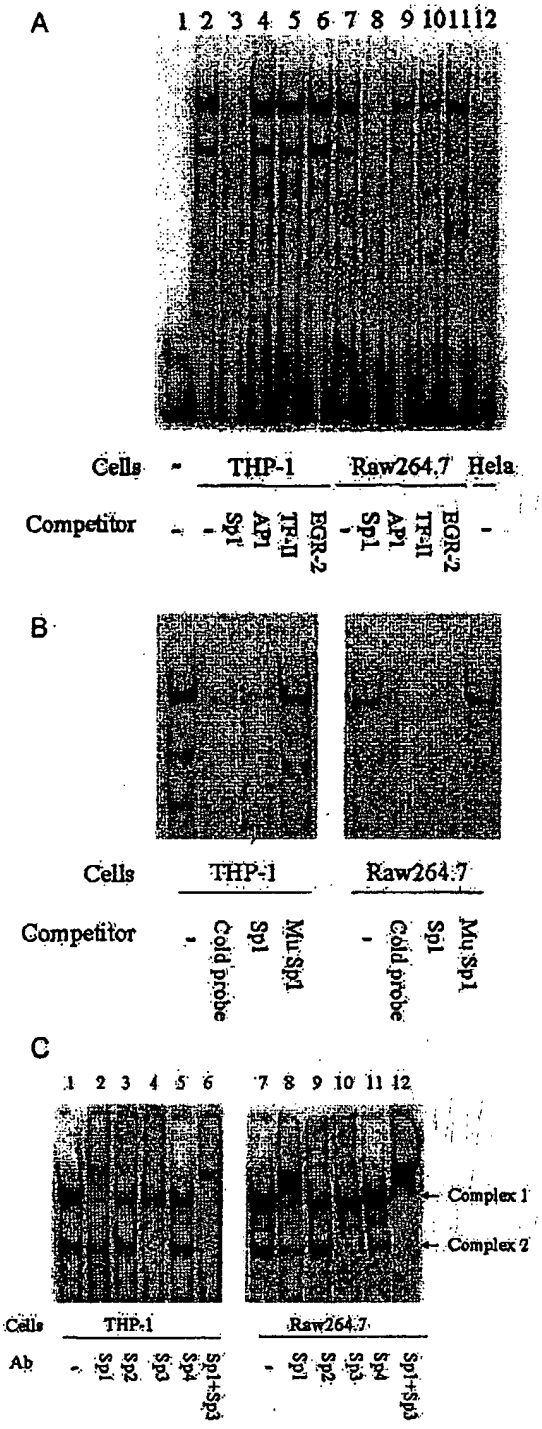


Fig. 4. Activity of deletions in the −471 bp region of the EMMPRIN promoter. Deletion constructs within −471 bp region of EMMPRIN promoter are shown in (A) and luciferase activities driven by these constructs in THP-1 cells (B) and Raw264.7 cells (C). Data shown as the percentage of maximal activity expressed as the mean ± SEM. *$P < 0.01$.

ATTCGATCGGTTCGGGGCGAG) was introduced into the competition EMSA experiment along with unlabeled wild-type 30 bp nucleotides (same sequence as labeled probe–cold probe) and Sp1 consensus oligonucleotides.

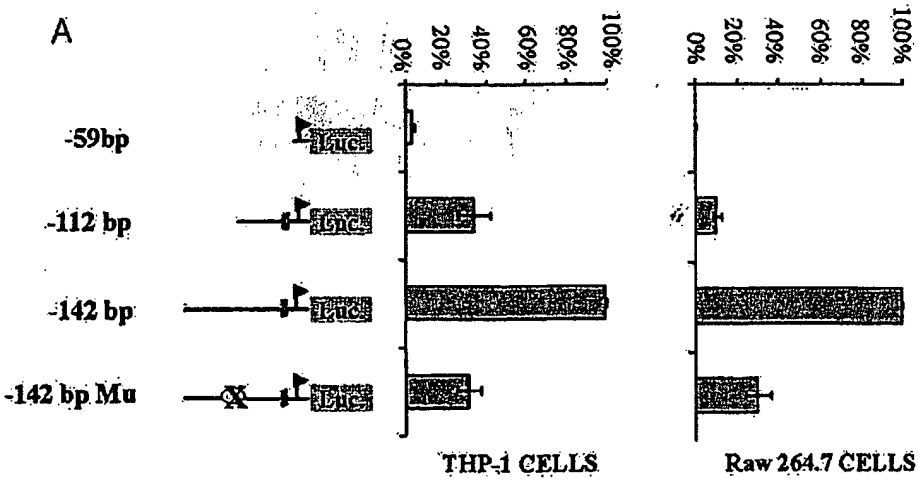


As illustrated in Fig. 5B, unlabeled mSp1 did not affect the formation of complexes while the unlabeled 30 bp wild-type oligonucleotide (cold probe) and Sp1 consensus oligonucleotide competed away the binding. To better define which proteins bind to the complex, we performed a supershift assay using antibodies against Sp1, Sp2, Sp3, and Sp4. As shown in Fig. 5C, complex 1 was completely supershifted in the presence of Sp1 antibody or partially shifted with the Sp3 antibodies. Complex 2 was completely shifted in the presence of Sp3 antibodies and partially shifted with the Sp1 antibody. Both complexes were shifted when both Sp1 and Sp3 antibodies were presented. No supershift was detected when Sp4 antibody was added. These results demonstrated that both Sp1 and Sp3 bind to the Sp1 DNA element in the EMMPRIN promoter.

3.5. Disruption of Sp1 consensus sequence on the proximal region of EMMPRIN promoter decreased promoter activity and eliminated DNA-protein complex formation

Within this 30 bp region, there is a Sp1 element located at −122 to −114 bp of the EMMPRIN promoter. To test that this element was important for EMMPRIN gene transcription in addition to binding with nuclear proteins, a site-directed substitution of the −122 to −116 bp region of DNA with AGATCTA was generated and subcloned into the −142 bp construct, since this construct was the shortest construct exhibiting high promoter activity. The luciferase activity driven by the native and mutated construct was measured. As show in Fig. 6A, mutation of the Sp1 element resulted in a 70% reduction of EMMPRIN promoter activity in both THP-1 and Raw264.7 cells. The promoter activity of this mutant construct was almost the same as that of the −112 bp construct in THP-1 cells. These results demonstrate that the Sp1 element in the −122 to −114 bp region of EMMPRIN promoter plays a functional role in promoting EMMPRIN gene transcription. The 30 bp oligonucleotide containing the Sp1 mutation was used as labeled probe to perform the EMSA with nuclear extracts of THP-1 and Raw264.7 macrophages. As shown in Fig. 6B, the Sp1 mutant probe significantly decreased the formation of DNA-protein complexes in both THP-1 and Raw264.7 cells (Fig. 6B, lanes 1 and 6) compared to that of wild-type probe (Fig. 6B, lanes 4 and 9). In addition, unlabeled

Fig. 5. Electrophoretic mobility shift assay of complex formation over the 30 bp region (−142 to −112 bp) of human EMMPRIN promoter. ^{32}P-Labeled 30 bp oligonucleotide (A, lane 1) was used as probe to form complex with nuclear proteins from human macrophages THP-1 (A, lanes 2–6) and mouse macrophages Raw264.7 (A, lanes 7–11) as well as HeLa cells (A, lane 12). Competition analysis was performed in the present of 100 fold molar excess of unlabeled Sp1, AP1, TF-II, EGR-2 consensus oligonucleotide (A) and unlabeled probe self (cold probe), Sp1 consensus oligonucleotide, mutant Sp1 consensus oligonucleotide (B). Supershift analysis was performed using antibody against Sp1, Sp2, Sp3, Sp4 and Sp1 plus Sp3 (C). Two major complexes that shifted in the presence of antibodies are indicated.



5'-CCGG CGTCCCCGGC GCTCGAGATCTACCCGAGAT GACGCCGTGC
GTGCGCGCGC CCGGTCCGCG CCTCCGCCGC TTTT TATAGC GGCCGCGGGC
GGCGGCGGCA GCGGTTGGAG TTGTAGGACC GGCGAGGAAT AGGAATCATG
GCGGCTGCGC TG

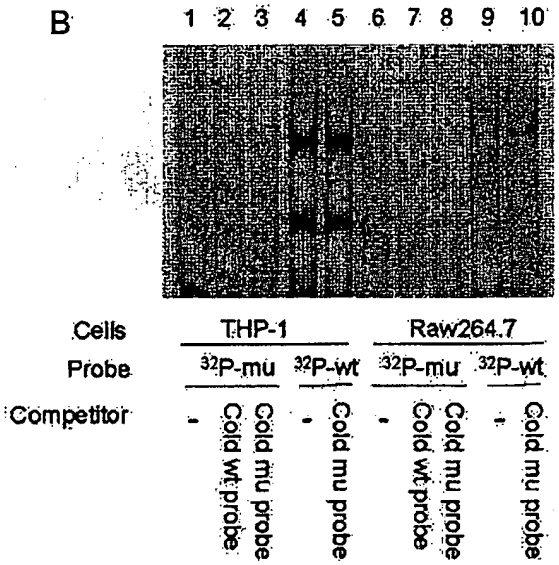


Fig. 6. Effect of site-specific mutation on human EMMPRIN promoter activity (A) and complex formation in the EMSA (B). Site-directed mutagenesis was performed to introduce a mutation located at −116 to −122 bp of −142 bp human EMMPRIN promoter construct. Sequence is shown with the mutation underline. The luciferase activities of mutant or the indicated constructs in THP-1 and Raw264.7 cells were measured and normalized against an internal control. The data are presented as percentage of maximal activity in mean ± SEM. In (B), ^{32}P-labeled mutant probe 5′-CCGGCGTCCCCGGCGCTCGAGATCTACCCG and wild-type (wt) probe were used to form complex with nuclear protein from THP-1, Raw264.7 and EMSA was performed with or without competition of 100-fold excess of unlabeled wt and mutant probe self (cold wt probe and cold mu probe).

Sp1 mutant oligonucleotide (cold mu probe) was unable to compete away the formation of complexes in this EMSA (Fig. 6B, lanes 5 and 10). These results indicate that the Sp1 element in this region specifically binds with Sp1 and Sp3 transcription factors.

3.6. Cooperative activation of human EMMPRIN promoter by Sp1 and Sp3

To determine how Sp1 and Sp3 functionally modulate the EMMPRIN promoter activity, Drosophila SL2 cells, which

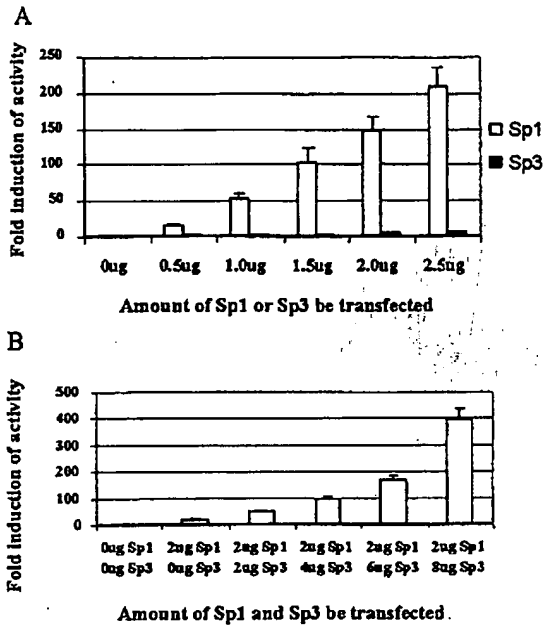


Fig. 7. Human EMMPRIN promoter activity in *Drosophila* SL2 cells. (A) Effect of Sp1 or Sp3 alone on EMMPRIN promoter. SL2 cells were cotransfected with 2.5 μg −142 bp human EMMPRIN promoter construct with various amount (0.5–2.5 μg) of pPacSp1 or pBKCMVSp3. Cells were harvested after 48 h incubation, and luciferase activity and the protein concentration were measured. The luciferase value was normalized by the protein concentration and fold induction represents the fold increase in the luciferase value of each sample against that of mock transfection. Data are shown as mean ± SEM. (B) Effect of Sp3 plus Sp1 on EMMPRIN promoter. SL2 cells were transfected with 2.5 μg −142 bp human EMMPRIN promoter construct with 2 μg of pPacSp1with various amounts (2–8 μg) of pBKCMVSp3 or salmon sperm DNA. Data are shown as mean ± SEM.

are deficient in Sp1 and Sp3, were utilized. SL2 cells were cotransfected with −142 bp EMMPRIN promoter construct and increasing concentration of Sp1 or Sp3 expression plasmids or both. As shown in Fig. 7, both Sp1 and Sp3 stimulated the −142 bp EMMPRIN promoter activity in a dose-dependent manner. Sp1 was the most potent transcription activator for the −142 bp EMMPRIN promoter (208-fold induction with 2.5 μg Sp1 plasmid transfection) while Sp3 was much less potent (8-fold induction with 2.5 μg Sp3 plasmid transfection). Cotransfection of 2 μg Sp1 and 2.5 μg of −142 EMMPRIN promoter construct with increasing amounts of the Sp3 plasmid (0–8 μg) resulted in a dose-dependent Sp3 enhanced activation, indicating that Sp1 and Sp3 activate the EMMPRIN promoter synergistically and Sp3 may act as an enhancer to augment the Sp1-mediated activation (Fig. 7B).

4. Discussion

Transcriptional regulation of the extracellular MMP inducer gene, EMMPRIN, in differentiated macrophages is shown in the present study to be dependent upon the Sp1 element within the −122 to −114 bp region of the EMMPRIN promoter. In addition, this functional Sp1 element is shown to be bound and positively regulated by both transcription factors Sp1 and Sp3. Several pieces of evidences support this conclusion: (1) deletion of the −1797 to −112 bp region of the EMMPRIN promoter significantly diminished the transcriptional activity compared to deletion of the −1797 to −142 bp region; (2) mutation of the Sp1 element (−122 to −116 bp) on the proximal region of the EMMPRIN promoter significantly decreased functional activity; (3) unlabeled Sp1 element could compete away the binding of EMMPRIN proximal promoter with nuclear protein; (4) both Sp1 and Sp3 formed complexes with EMMPRIN promoter DNA, which was recognized by antibodies against both proteins in a supershift EMSA; (5) in the Sp1 and Sp3 deficient cells SL-2, −142 bp EMMPRIN promoter did not exhibit any luciferase activity; (6) transfection of Sp1 or Sp3 or both into SL-2 insect cells increased the EMMPRIN transcription activity in a dose-dependent manner.

The EMMPRIN 5′-flanking region (950 bases) and the transcription start site were previously reported by Guo et al. (1998). Sequence alignment of the human and mouse EMMPRIN promoters, show that three Sp1 and two AP2 putative transcriptional binding sites are highly conserved within the mouse EMMPRIN promoter, implicating these Sp1 or AP2 sites as regulatory elements for EMMPRIN gene expression. Consistent with this inference, we mapped the EMMPRIN promoter by deletion analysis and revealed that the proximal region (−142 to −112 bp) is essential for human EMMPRIN promoter activity in macrophages. In addition, we identified that a Sp1 element within this region, which is conserved within human and mouse EMMPRIN promoter is essential for EMMPRIN promoter activity by site-directed mutagenesis, which resulted in 100% (THP-1 cells) and 78% (Raw264.7 cells) loss of promoter activity, respectively. Further, supershift analysis revealed that both Sp1 and Sp3 bind to this functional 30 bp oligonucleotide. Sp1 is a well-characterized transcription factor that binds to the Sp1 consensus element present in a variety of cellular and viral promoters and stimulate their transcriptional activity (Suske, 1999). It is also known that there are a number of Sp1 element-binding factors homologous to Sp1, named Sp2, Sp3 and Sp4. Sp3 and Sp4 have the most conserved amino sequence of the DNA binding domain and recognize the GC-rich Sp1 binding element with similar specificity and affinity as Sp1. Sp2 does not bind to the classical GC-rich Sp1 element but to the GT-rich element (Suske, 1999). Sp1 has been implicated in the activation of a large number of genes and it is thought to be involved in several cellular processes; but less is known about Sp3. Sp3 contains two glutamine-rich domains that have a strong activation function and a transcriptional repression domain that can silence the activation modules. (Dennig et al., 1996). Originally,

Sp3 was found to suppress Sp1 mediated activation by binding to the same site and thereby preventing Sp1 binding and activation (Hagen et al., 1994; Majello et al., 1997). However, this model has been challenged by the observation of stimulatory transcriptional activity for Sp3 (Majello et al., 1997). This later observation is consistent with our findings showing that both Sp1 and Sp3 transactivate EMMPRIN promoter in *Drosophila* SL2 cells, even though Sp3 had less of an effect. Of most interest is that Sp1 and Sp3 function in a cooperative manner for EMMPRIN promoter activity (Fig. 7B). However, the molecular mechanism for the activation function of Sp3 and the reason for the observed synergistic action of Sp1 and Sp3 on EMMPRIN promoter activity are unclear.

Although Sp1 and Sp3 are ubiquitous nuclear factors, their expression levels vary greatly during different stages of cellular development or in different cell types (Saffer et al., 1991). It has been reported that the Sp1 protein is markedly increased in PMA-induced differentiated THP-1 cells and this protein is responsible for induction of acid sphingomyelinase expression during THP-1 differentiation (Langmann et al., 1999). We have compared the Sp1 and Sp3 protein levels in human monocytes versus human monocyte derived macrophage by Western analysis of nuclear extracts and found that both Sp1 and Sp3 were upregulated during differentiation (Fig. 8). These results raised the possibility that higher expression of Sp1 and Sp3 in the macrophage than in monocytes not only account for higher basal expression of EMMPRIN, but also account for upregulation of EMMPRIN expression during monocyte differentiation.

This ubiquitous expression of Sp1 and Sp3 may account for the expression of EMMPRIN in various cells (i.e. macrophage, cancer cells) and tissues. However, the binding and activation of Sp1 and Sp3 are regulated by variety of stimuli, which may account for regulation of gene expression, cellular growth and differentiation. These biological stimuli modulate Sp through post-transcriptional modification, such as glycosylation and phosphorylation (Armstrong et al., 1997; Jackson and Tjian, 1988). Other regulatory mechanisms, such as Sp1 and Sp3 interacting with other transcription factors and synergistically activating transcription, have been reported (Galvagni et al., 2001; Yieh et al., 1995). In the present study, we have found that Sp3 enhance Sp1-mediated EMMPRIN promoter activity in SL2 cells, even though Sp3 alone has minimal transcriptional activity. Based on this finding we speculate that Sp3 may regulate the transcription activity via cooperation with Sp1. In addition, we have mapped a negative regulatory region of EMMPRIN promoter. Whether this region has any functional implication for Sp1 and Sp3 activation which leads to EMMPRIN gene suppression still needs further investigation.

In summary, we have provided basic information relating to how the human EMMPRIN gene is regulated in macrophages. Given the abundance of macrophages within atherosclerotic lesions and the known function of EMMPRIN to induce MMPs, future investigation into the role of macrophage EMMPRIN expression on vascular cells infiltration and migration are warranted.

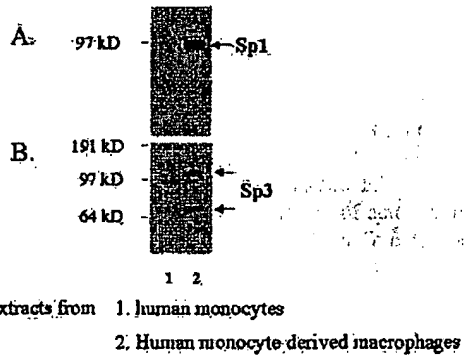


Fig. 8. Sp1 (A) and Sp3 (B) levels in nuclear extracts from human monocytes and differentiated macrophages. Twenty micrograms of nuclear extracts from human monocytes or human monocyte-derived macrophages were loaded on 4–12% NuPAGE gel and immunoblotted with polyclonal antibodies against Sp1 (A) and Sp3 (B). Molecular mass marker is indicated at left.

Acknowledgements

We thank Dr. Margaret Johns for genome and putative transcription factor search, Kenneth Hoppe and Dr. Omar Francone for pPacSp1, pBKCMVSp3 plasmid and SL2 cells, Jing Chen for total RNA of mouse kidney, Dr. Sun Yi, Dr. Sotirios Karathanasis, Dr. Mark Rekhter for helpful discussions, and Dr. Tom Hullinger, Marina Kvitnaya, Zhiwu Lin, Xiangyang Xu, Brian Batley, and Patrick Gillespie for technical suggestions.

References

Amorino, G.P., Hoover, R.L., 1998. Interactions of monocytic cells with human endothelial cells stimulate monocytic metalloproteinase production. Am. J. Pathol. 152, 199–207.

Armstrong, S.A., Barry, D.A., Leggett, R.W., Mueller, C.R., 1997. Casein kinase II-mediated phosphorylation of the C terminus of Sp1 decreases its DNA binding activity. J. Biol. Chem. 272, 13489–13495.

Biswas, C., Zhang, Y., DeCastro, R., Guo, H., Nakamura, T., Kataoka, H., Nabeshima, K., 1995. The human tumor cell-derived collagenase stimulatory factor (renamed EMMPRIN) is a member of the immunoglobulin superfamily. Cancer Res. 55, 434–439.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156–159.

Dennig, J., Beato, M., Suske, G., 1996. An inhibitor domain in Sp3 regulates its glutamine-rich activation domains. EMBO J. 15, 5659–5667.

Galis, Z.S., Sukhova, G.K., Lark, M.W., Libby, P., 1994. Increased expression of matrix metalloproteinases and matrix degrading activity in vulnerable regions of human atherosclerotic plaques. J. Clin. Invest. 94, 2493–2503.

Galis, Z.S., Sukhova, G.K., Kranzhofer, R., Clark, S., Libby, P., 1995. Macrophage foam cells from experimental atheroma constitutively

produce matrix-degrading proteinases. Proc. Natl. Acad. Sci. USA 92, 402–406.

Galvagni, F., Capo, S., Oliviero, S., 2001. Sp1 and Sp3 physically interact and co-operate with GABP for the activation of the utrophin promoter. J. Mol. Biol. 306, 985–996.

Guo, H., Zucker, S., Gordon, M.K., Toole, B.P., Biswas, C., 1997. Stimulation of matrix metalloproteinase production by recombinant extracellular matrix metalloproteinase inducer from transfected Chinese hamster ovary cells. J. Biol. Chem. 272, 24–27.

Guo, H., Majmudar, G., Jensen, T.C., Biswas, C., Toole, B.P., Gordon, M.K., 1998. Characterization of the gene for human EMMPRIN, a tumor cell surface inducer of matrix metalloproteinases. Gene 220, 99–108.

Hagen, G., Muller, S., Beato, M., Suske, G., 1994. Sp1-mediated transcriptional activation is repressed by Sp3. EMBO J. 13, 3843–3851.

Jackson, S.P., Tjian, R., 1988. O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. Cell 55, 125–133.

Joris, I., Zand, T., Nunnari, J.J., Krolikowski, F.J., Majno, G., 1983. Studies on the pathogenesis of atherosclerosis. I. Adhesion and emigration of mononuclear cells in the aorta of hypercholesterolemic rats. Am. J. Pathol. 113, 341–358.

Kataoka, H., DeCastro, R., Zucker, S., Biswas, C., 1993. Tumor cell-derived collagenase-stimulatory factor increases expression of interstitial collagenase, stromelysin, and 72-kDa gelatinase. Cancer Res. 53, 3154–3158.

Langmann, T., Buechler, C., Ries, S., Schaeffler, A., Aslanidis, C., Schuierer, M., Weiler, M., Sandhoff, K., de Jong, P.J., Schmitz, G., 1999. Transcription factors Sp1 and AP-2 mediate induction of acid sphingomyelinase during monocytic differentiation. J. Lipid Res. 40, 870–880.

Lendon, C.L., Davies, M.J., Born, G.V., Richardson, P.D., 1991. Atherosclerotic plaque caps are locally weakened when macrophages density is increased. Atherosclerosis 87, 87–90.

Liao, H.S., Kodama, T., Doi, T., Emi, M., Asaoka, H., Itakura, H., Matsumoto, A., 1997. Novel elements located at −504 to −399 bp of the promoter region regulated the expression of the human macrophage scavenger receptor gene in murine macrophages. J. Lipid Res. 38, 1433–1444.

Libby, P., Clinton, S.K., 1993. The role of macrophages in atherogenesis. Curr. Opin. Lipidol. 4, 355–363.

Majello, B., De Luca, P., Lania, L., 1997. Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. J. Biol. Chem. 272, 4021–4026.

Major, T., Lu, X., Dagle, C., Bocan, T., 2000. EMMPRIN (extracellular matrix metalloproteinase inducer), a human tumor-cell-derived protein, is induced upon monocyte differentiation and is expressed in human atheroma. Circulation 102 (Suppl. II), II39.

Prescott, J., Troccoli, N., Biswas, C., 1989. Coordinate increase in collagenase mRNA and enzyme levels in human fibroblasts treated with the tumor cell factor, TCSF. Biochem. Int. 19, 257–266.

Ross, R., 1999. Atherosclerosis – an inflammatory disease. N. Engl. J. Med. 340, 115–126.

Saffer, J.D., Jackson, S.P., Annarella, M.B., 1991. Developmental expression of Sp1 in the mouse. Mol. Cell Biol. 11, 2189–2199.

Sameshima, T., Nabeshima, K., Toole, B.P., Yokogami, K., Okada, Y., Goya, T., Koono, M., Wakisaka, S., 2000. Glioma cell extracellular matrix metalloproteinase inducer (EMMPRIN) (CD147) stimulates production of membrane-type matrix metalloproteinases and activated gelatinase A in co-cultures with brain-derived fibroblasts. Cancer Lett. 157, 177–184.

Sun, J., Hemler, M.E., 2001. Regulation of MMP-1 and MMP-2 production through CD147/extracellular matrix metalloproteinase inducer interactions. Cancer Res. 61, 2276–2281.

Suske, G., 1999. The Sp-family of transcription factors. Gene 238, 291–300.

Yieh, L., Sanchez, H.B., Osborne, T.F., 1995. Domains of transcription factor Sp1 required for synergistic activation with sterol regulatory element binding protein 1 of low density lipoprotein receptor promoter. Proc. Natl. Acad. Sci. USA 92, 6102–6106.

Published by Oxford University Press *Human Molecular Genetics, 2001, Vol. 10, No. 26 3101–3109*

Effect of allelic variation at the NACP–Rep1 repeat upstream of the α-synuclein gene (*SNCA*) on transcription in a cell culture luciferase reporter system

Omit Chiba-Falek and Robert L. Nussbaum*

Genetic Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, 49 Convent Drive, MSC 4472, Bethesda, MD 20892-4472, USA

Received October 15, 2001; Revised and Accepted October 26, 2001

Mutations in the α-synuclein gene (*SNCA*) have been implicated in familial Parkinson's disease (PD) while certain polymorphic alleles at a microsatellite repeat, NACP–Rep1, located ~10 kb upstream of the gene, have been associated with sporadic PD. In order to study the regulation of the human α-synuclein gene, we performed a deletion analysis of 10.7 kb upstream of the translational start site, using the luciferase reporter assay in 293T cells and the neuroblastoma cell line SH-SY5Y. The shortest fragment, 400 bp upstream of the transcriptional start site, was sufficient for transcription in both cell lines. The other constructs led to variable expression levels, with some showing maximum expression and others showing nearly complete extinction of expression. An 880 bp fragment located ~10 kb upstream of the gene and containing the NACP–Rep1 polymorphism, was shown to be necessary for normal expression. Additional analysis of the NACP–Rep1 locus and surrounding DNA suggested that two domains flanking the repeat interact to enhance expression while the repeat acts as a negative modulator. Next, we measured the activity of the entire 10.7 kb upstream region in the luciferase reporter assay when each of our different NACP–Rep1 alleles were present. The expression levels varied very significantly among the different alleles over a 3-fold range in the SH-SY5Y cells but showed little or no significant variation in the 293T cells. Given that even small changes in α-synuclein expression may, over many decades, predispose to PD, the association of different NACP–Rep1 alleles with PD may be a consequence of polymorphic differences in transcriptional regulation of α-synuclein expression resulting from different NACP–Rep1 alleles.

INTRODUCTION

Parkinson's disease (PD) is the second most common neurodegenerative disease in humans; its etiology is largely unknown. The α-synuclein gene (*SNCA*) was first identified as the gene encoding a protein of which a subfragment, termed the non-β-amyloid component, was thought to be a component of Alzheimer's disease plaques (1). Interest in α-synuclein increased substantially when early onset PD, which can occur as an autosomal dominant trait in a few rare families, was shown to result from two different missense mutations in *SNCA* (2,3). The importance of α-synuclein in PD was further underscored by the demonstration of α-synuclein in the characteristic protein aggregates, termed Lewy bodies, found in the affected portions of the brains of sporadic PD patients (4). Recently, it was shown that both mice and flies expressing the human α-synuclein recapitulate some characteristics of PD (5,6). These findings suggest that α-synuclein could have an important role in the development of PD. However, efforts to identify *SNCA* mutations in sporadic PD have failed (7).

NACP–Rep1 is a polymorphic complex repeat site located ~10 kb upstream of the translational start of *SNCA* (8,9). Five alleles were identified with a size difference of two nucleotides (8,10). The basis for the size difference among the different NACP–Rep1 alleles was unknown. Recently, it was shown that certain alleles of the NACP–Rep1 locus are associated with an increased risk of sporadic PD in German and American populations (11–13) but not in the Japanese population (14). Variation in expression of the gene might play a role in the pathogenesis of the disease in some patients that do not carry the mutated protein. Hence, understanding the regulation of the α-synuclein expression could be very important.

In this study we aimed to characterize the α-synuclein promoter/enhancer region and to further study the role of the various NACP–Rep1 polymorphic alleles in regulating the expression of the *SNCA* gene.

RESULTS

Computational analysis of the α-synuclein putative promoter/enhancer region

We searched for potential transcription binding sites in the sequence upstream of *SNCA* transcriptional start using the MatInspector V2.2 software. The analysis revealed numerous potential recognition sites for many transcription factors (data not shown). The positions of the binding sites for AP1, C/EBPβ,

*To whom correspondence should be addressed. Tel: +1 301 402 2039; Fax: +1 301 402 2170; Email: rlnuss@nhgri.nih.gov

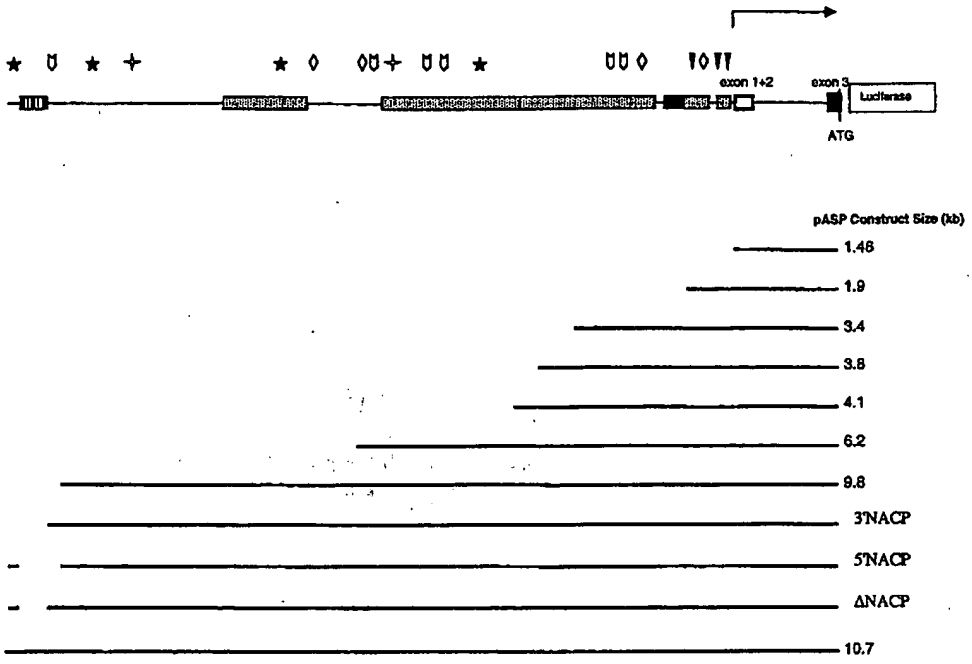


Figure 1. Schematic representation of the 10.7 kb region upstream of the α-synuclein gene with serial deletion constructs as shown. Boxes show areas of similarity to mouse sequence. The closed boxes represent areas of homology to mouse. The box with vertical lines represents the NACP–Rep1 repeat. The open box is exons 1 + 2 which showed poor similarity to mouse sequence. The arrow above indicates the transcription start site (accession no. U46896). The translational start site is marked in exon 3 by ATG. Below are the different constructs. The size (kb) from the translational start site of each construct is indicated at its right. The positions of the predicted binding sites for the following transcription factors are marked above: open stars, AP1; open arrowheads, C/EBPβ; closed stars, CHOP (GADD 153); diamonds, N-Myc; closed arrowheads, Sp1.

CHOP (GADD 153), N-Myc and Sp1, which gave high score values (core similarity = 1; matrix similarity ≥ 0.85), are shown in Figure 1.

Analysis of the deleted constructs at the promoter/enhancer region of α-synuclein

The 10.7 kb region upstream of *SNCA* translational start (9) was used to make eight constructs extending varying distances upstream of the transcriptional start site; all constructs were inserted in a luciferase expression vector and transfected into 293T cells and the neuroblastoma cell line SH-SY5Y. Each plasmid was cotransfected with pRL-TK (293T cells) or pRL-SV40 (SH-SY5Y cells) and the firefly and Renilla luciferases expression were measured. As a control for the luciferase basal expression we cotransfected the commercial promoter-less plasmid, pGL 3-Basic, with the pRL plasmid. For each cotransfection experiment, the relative activity of luciferase was calculated in order to eliminate the effect of transfection efficiency and cell number (Materials and Methods).

The different constructs harboring sequences upstream of the transcriptional start site were assayed for expression of luciferase. The control plasmid, pAS-1.46 which includes only sequences downstream of the transcriptional start site gave the lowest luciferase activity in both cell lines, only a few fold over the background activity seen with the empty luciferase vector. The addition of the immediate 400 bp upstream of the transcription start site increased the luciferase expression 5- and 8-fold in 293T and SH-SY5Y cells, respectively. Thus, the shortest fragment, 400 bp upstream of the transcriptional start site, was found to be sufficient for transcription in both cell lines (Fig. 2) and is likely to include the minimal promoter. The other constructs led to variable expression levels, with some showing maximum expression and others showing nearly complete extinction of expression (Fig. 2). As can be seen in Figure 2 there were also differences in relative expression level for each construct between the cell lines. These results illustrate that this genomic region contains promoter/enhancer elements that are likely to belong to the *SNCA* gene and might contribute to tissue specific regulation of *SNCA* transcription.

Our previous studies showed the importance of the region containing the human NACP–Rep1 repeat in the regulation of expression of the human α-synuclein gene (9). The construct harboring the NACP–Rep1 repeat region at the 5′ end (pASP-10.7) resulted in a 60- and 3-fold increase in luciferase expression relative to the basal level of pGL 3-Basic upon transfection into 293T and SH-SY5Y cells, respectively (Fig. 2). The construct in which the NACP–Rep1 repeat region was deleted (pASP-9.8) led to a decrease in the expression level to 15-fold in 293T relative to the basal level of pGL 3-Basic (Fig. 2A). Thus, the 880 bp segment containing NACP–Rep1 contributes a 4-fold increase in the α-synuclein promoter activity in 293T cells. The relative expression level of pASP-9.8 in SH-SY5Y

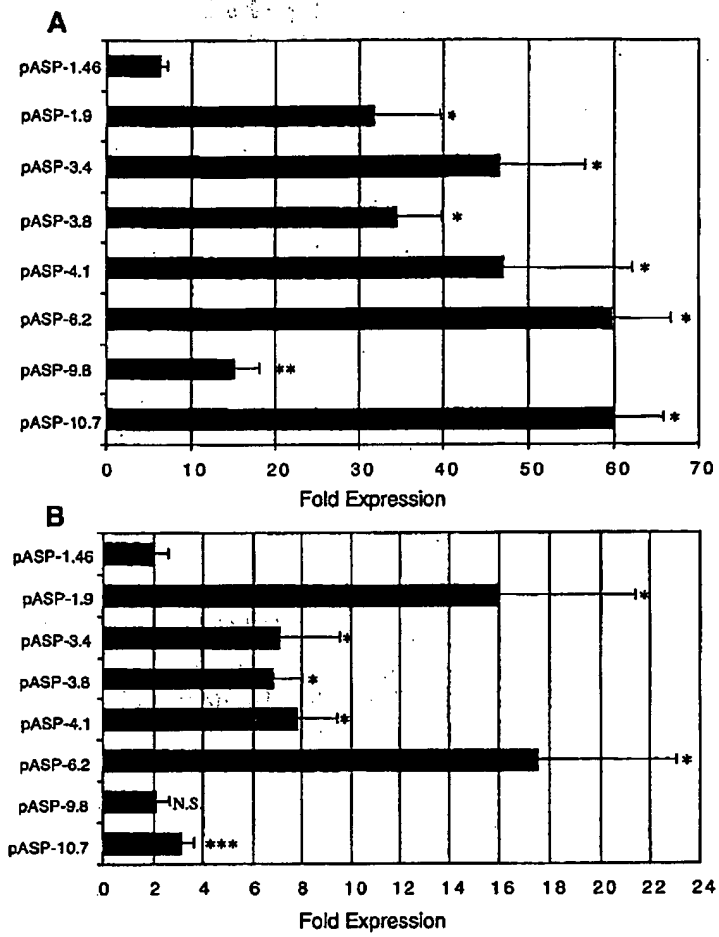


Figure 2. The fold expression of luciferase activity demonstrated by the full-length and the seven serial deletion constructs of pASP in 293T (A) and SH-SY5Y (B) cells. Cells were transfected with each construct or pGL 3-Basic; in the case of pASP-10.7, activity shown is with the construct carrying the allele with the lowest expression (allele 0). Data shown here are the means ± 1 SEM of three to six independent experiments performed on separate days. The relative activity with each pASP or pGL 3-Basic plasmid was calculated by dividing the luminescence intensity of the firefly luciferase by that of the cotransfected Renilla luciferase in each independent aliquot of cells and then averaging the three relative luciferase activities seen. The fold expression for each pASP was then determined by dividing the average relative activity of each construct to that of the average obtained with pGL 3-Basic. The differences in expression seen with the different constructs was highly significant ($P << 10^{-15}$, ANOVA). *P*-values for two-tailed *t*-test comparing each pASP construct to pASP-1.46 are: $*P < 5 \times 10^{-5}$; $**P = 2 \times 10^{-4}$; $***P < 0.02$; N.S., not significant.

cells was similar to the level observed from the control plasmid, pAS-1.46 in which no promoter sequences were included (Fig. 2B). Thus, deletion of the −880 bp region containing the NACP–Rep1 repeat led to the complete elimination of the expression derived by the *SNCA* promoter in SH-SY5Y cells. However, of interest was that in both cell lines, deletion of an additional 3.6 kb of sequence from the 5′ end of the pASP-9.8 construct to generate the pASP-6.2 construct caused a return to the highest levels seen.

We further studied the ~880 bp region containing the NACP–Rep1 repeat. A construct carrying a precise deletion of only the NACP–Rep1 (pASP-ΔNACP) led to a 7- and 3.5-fold increase in expression relative to the 880 bp deletion in 293T and SH-SY5Y cells, respectively (Fig. 3). In 293T cells deletion of the NACP–Rep1 plus 317 bp upstream (pASP-3′NACP) or the NACP–Rep1 plus 452 bp downstream (pASP-5′NACP) of the repeat led to 5.5- or 3.8-fold increase in expression compared to the 880 bp deletion (Fig. 3A). Similar results were shown in SH-SY5Y cells: deletion of the NACP–Rep1 and 317 bp upstream (pASP-3′NACP) or 452 bp downstream (pASP-5′NACP) of the repeat led to 2.3- or 1.6-fold increase in expression compared to the 880 bp deletion (Fig. 3B).

Sequence of the different alleles of the complex NACP–Rep1 repeat

According to sequences deposited in GenBank, human NACP–Rep1 alleles are composed of the following dinucleotides: (TC)(TT)(TC)(TA)(CA) with variable numbers of the

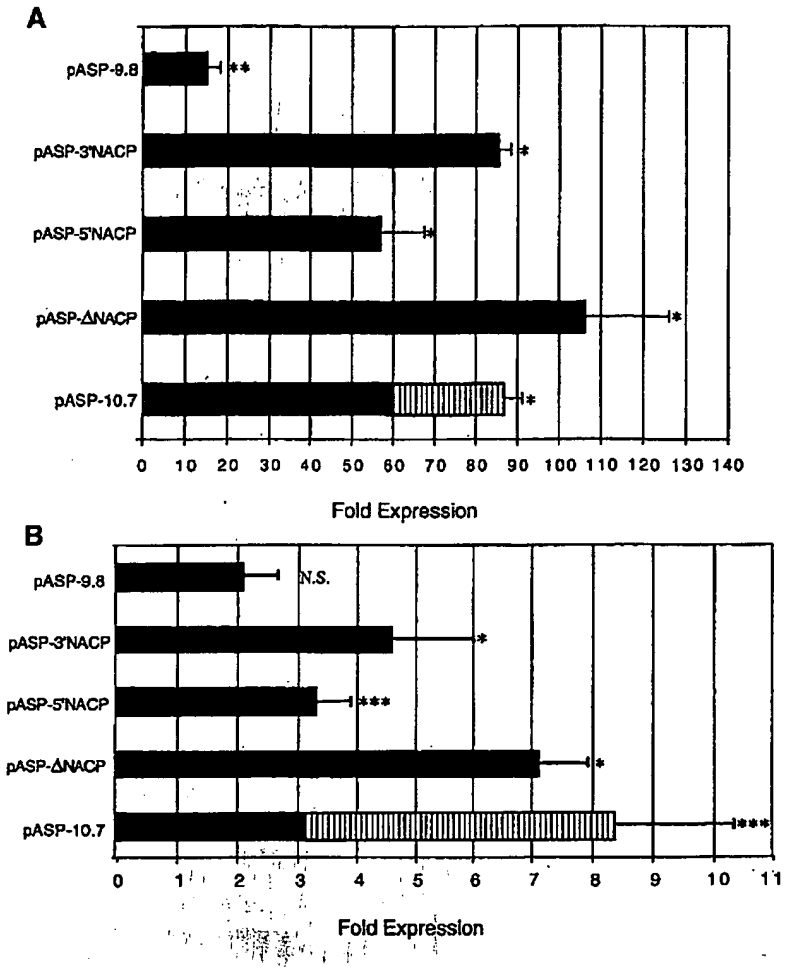


Figure 3. The fold expression of luciferase activity derived by different pASP constructs carrying deletions of the NACP–Rep1 site and its flanking region in 293T (A) and SH-SY5Y (B) cells. Cells were transfected with each construct or pGL 3-Basic. The relative activity with each pASP or pGL 3-Basic plasmid was calculated by dividing the luminescence intensity of the firefly luciferase by that of the cotransfected Renilla luciferase in each independent aliquot of cells and then averaging the three relative luciferase activities seen. The fold expression for each deleted pASP was then determined by dividing the average relative activity of each construct to that of the average obtained with pGL 3-Basic. Data shown here are the means ± 1 SEM of four to eight independent experiments. The hatched portion of the pASP-10.7 bar covers the expression range of all four NACP–Rep1 alleles (Fig. 4). *P*-values for two-tailed *t*-test comparing each pASP construct to pASP-1.46 are: $*P < 5 \times 10^{-5}$; $**P = 2 \times 10^{-4}$; $***P < 0.02$; N.S., not significant.

TC, TA and CA dinucleotide repeats; the sequences of alleles 0 and 1 were reported as $(TC)_{10}(TT)_1(TC)_{10}(TA)_8(CA)_{10}$ and $(TC)_{10}(TT)_1(TC)_{10}(TA)_8(CA)_{11}$, respectively (accession nos AC015529, AC022357 and AP001947). In order to explore more fully the sequence differences among the different NACP–Rep1 alleles, we screened for individuals carrying alleles designated 0, 1, 2 and 3 based on the size of the alleles (8,10). Next, PCR products from individuals homozygous for alleles 0, 1 or 2 (eight, 12 and two chromosomes, respectively) and from an individual heterozygous for alleles 1 and 3, were subcloned into the pCR-TOPO vector and subsequently sequenced. We approached the problem of defining the actual sequence of these various alleles by sequencing independent, cloned PCR products with the assumption that sequences arising from PCR artifacts will be seen infrequently among the cloned products while the correct allele sequence will be seen repeatedly in the cloned PCR products. Using this approach, we compiled the sequences from the various alleles and observed that variation in the CA element is most likely responsible for the different length alleles. We confirmed that human NACP–Rep1 alleles are composed of the following dinucleotides: (TC)(TT)(TC)(TA)(CA) with variable numbers of the TC, TA and CA dinucleotide repeats (Table 1). Our analysis of all this sequence data suggested that the size variation among the different alleles is mainly due to the numbers of CA repeats, with 10, 11, 12 and 13 repeats present in alleles 0, 1, 2

Table 1. Sequences of the different NACP–Rep1 alleles

	TC	T	TC	TA	CA	No. of clones
0/0						
A	10	2	9	9	10	1/5
	10	2	10	8	10	4/5
B	10	2	10	8	10	3/5
	10	2	10	7	11	2/5
C	10	2	11	7	10	1/4
	10	2	10	8	10	3/4
D	10	2	11	7	10	5/9
	10	2	9	9	10	4/9
Most frequent	10	2	10	8	10	10/23
1/1						
E	11	2	11	7	10	1/4
	10	2	10	8	11	3/4
F	11	2	11	7	10	2/4
	10	2	10	8	11	1/4
	11	2	10	8	10	1/4
G	11	2	10	8	10	2/5
	10	2	10	8	11	3/5
H	10	2	10	8	11	2/5
	11	2	9	9	10	2/5
	10	2	9	9	11	1/5
I	11	2	10	8	10	1/5
	10	2	10	8	11	3/5
	10	2	9	9	11	1/5
J	10	2	10	8	11	1/6
	10	2	11	7	11	5/6
Most frequent	10	2	10	8	11	13/29
2/2						
K	10	2	10	8	12	6/7
	10	2	9	9	12	1/7
Most frequent	10	2	10	8	12	6/7
Het 3						
L	10	2	10	8	13	3/5
	10	2	9	9	13	1/5
	10	2	10	9	12	1/5
Most frequent	10	2	10	8	13	3/5

The sequence of the NACP–Rep1 alleles was analyzed in individuals homozygous for allele 0 (A–D), allele 1 (E–J), allele 2 (K) and in an individual heterozygous for allele 3 (L) (Materials and Methods). For each individual all the sequence combinations for a certain allele size are presented. The number of clones for each sequence combination from the total number of clones analyzed in a specific individual is indicated in the 'No. of clones' column. For each genotype the most frequent sequence combination is presented in the row designated 'Most frequent'. The number of clones carrying the most frequent sequence out of the total number of clones analyzed in all individuals carrying a specific genotype is indicated under 'No. of clones'.

and 3, respectively. Interestingly, alleles 1 and 0 may represent more than a single allele each since some individuals apparently homozygous for each of these alleles based on the length of the amplified product seemed to carry more than one allele at the sequence level (Table 1). However, these results could not be demonstrated without the use of PCR and it cannot be excluded that these results arise from PCR errors.

Analysis of the expression derived by constructs carrying different NACP–Rep1 alleles

To determine whether the variable length of the complex repeat is able to modulate the transcription of the α-synuclein gene, an 880 bp region of the native *SNCA* promoter/enhancer encompassing the NACP–Rep1 site was amplified from individuals carrying each of the alleles and cloned into the ASP-10.7 by replacing its original NACP–Rep1 allele (allele number –1). We further studied the promoter/enhancer activity of the NACP–Rep1 element in the neuroblastoma cells SH-SY5Y since it is more representative of the *in vivo* tissue in which α-synuclein is expressed. Each construct was cotransfected with pRL-SV40 into SH-SY5Y cells and the expression compared to that seen with the promoter-less plasmid, pGL 3-Basic. The different constructs harboring the variable length CA repeat led to different levels of luciferase expression (Fig. 4A). In SH-SY5Y cells, the construct containing the largest repeat (13 CA) resulted in a 2.5-fold increase in activity over that seen with the pASP-10.7, which contains 10 CA repeats. It is of interest to note that this allele had been shown to be associated with sporadic PD in a German and an American population (10,12,13). The promoter activity was suppressed as the length of the repeat decreased to 12 CA, resulting in only a 1.5-fold increase in activity as compared to the shortest repeat. The highest promoter activity was seen with the construct containing the repeat with 11 CA, leading to a 3-fold increase in the expression relative to pASP-10.7 (Fig. 4A). In contrast, there were only minor changes of borderline statistical significance in expression levels seen in 293T cells among the different constructs harboring the variable length of CA repeat (Fig. 4B).

DISCUSSION

The vast majority of PD patients do not show Mendelian inheritance of their disease and have no mutations in α-synuclein, yet they have deposits of α-synuclein in pathognomonic aggregates, Lewy bodies and Lewy neurites, in the region of the substantia nigra as well as in other locations in the brain (15). One of the more important outstanding questions for research into the molecular pathogenesis of PD is the role of α-synuclein in sporadic PD. The involvement of wild-type α-synuclein in PD could be the result of post-translational modification or damage to the protein, altered regulation of expression, abnormal degradation or some combination of all three. In this study we explore the role of the region upstream of the *SNCA* gene in regulation of α-synuclein expression and further show the contribution of the polymorphic NACP–Rep1 element to modulating the activity of the α-synuclein promoter/enhancer.

First, we examined the region upstream of the *SNCA* gene for its basal promoter. A comparison of the luciferase reporter gene activity in pASP-1.9 and pASP-1.46 revealed that the –400 bp DNA fragment located immediately 5′ of the reported transcriptional start site is sufficient to drive expression. Computational analysis revealed three potential Sp1 binding sites, but no TATA box, in this –400 bp segment. These

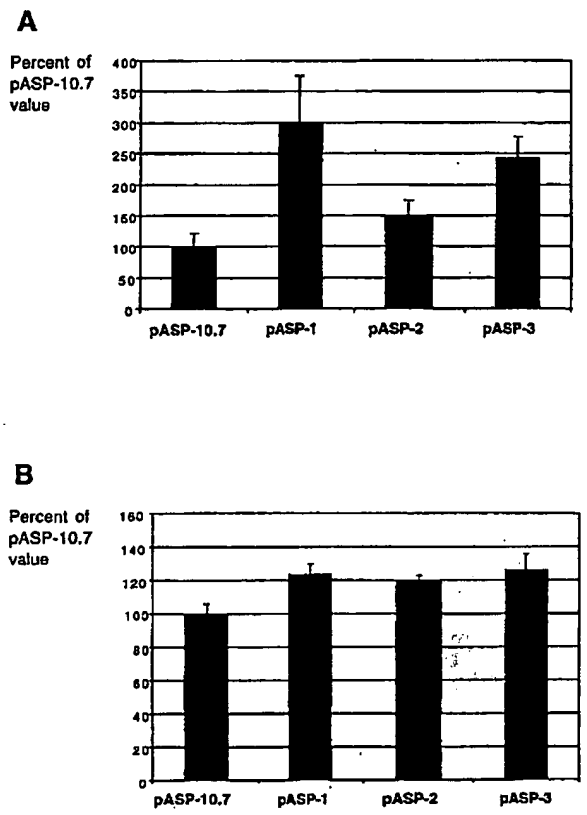


Figure 4. The ratio of fold expression of luciferase activity derived by the full-length pASP constructs harboring different NACP–Rep1 alleles to pASP-10.7 in SH-SY5Y (A) and 293T (B) cells. Cells were cotransfected with each of the four constructs including different NACP–Rep1 alleles or pGL 3-Basic and pRL. For each construct four to five experiments were performed. The relative activity with each pASP or pGL 3-Basic plasmid was calculated by dividing the luminescence intensity of the firefly luciferase by that of the cotransfected Renilla luciferase in each independent aliquot of cells and then averaging the three relative luciferase activities seen. The fold expression for each pASP was then determined by dividing the average relative activity of each construct to that of the average obtained with pGL 3-Basic. Then, the ratio in percentage of the fold expression for each of the pASP-1, pASP-2 and pASP-3 relative to pASP-10.7 was determined. The average of the ratios of four to five independent experiments performed on separate days was calculated. The fold expression of pASP-10.7 is arbitrarily assigned 100%. The data represented here are the average ratios in percentage ± 1 SEM for the pASP-1, pASP-2 and pASP-3 constructs relative to pASP-10.7. Student's *t*-test comparing the 'fold expression' of each of the pASP-1, pASP-2 and pASP-3 constructs to pASP-10.7 revealed $P = 5 \times 10^{-6}$, 0.002 and 5×10^{-9}, respectively, in SH-SY5Y cells, and $P = 0.06$, 0.08 and 0.04, respectively, in 293T cells.

features resemble previously described promoters of neuron-specific genes; such as Tau (16,17), the α5 subunit of the neuronal nicotinic receptor (18), the *N*-methyl-D-aspartate receptor subunit (NMDAR1) (19) and the recently characterized promoter of the human parkin gene (20). This is also in the size range of many common basal promoters but does not rule out that the basal promoter might be even shorter than the smallest fragment we studied.

Sequence analysis of the rest of the 10.7 kb DNA fragment upstream of *SNCA* predicted numerous additional potential binding sites for a number of transcription factors. An examination of promoter/enhancer activity of deletion constructs involving the upstream 10.7 kb sequence of the *SNCA* gene confirmed the importance of this region in control of α-synuclein expression. In both 293T and SH-SY5Y cells, pASP-6.2 showed the highest levels of expression whereas pASP-9.8 had markedly lower expression levels, suggesting that perhaps a silencer element may reside within the 3.6 kb region included in pASP-9.8 but missing from pASP-6.2. The 880 bp segment present in pASP-10.7 and absent in pASP-9.8 can overcome, at least in part, the negative effect on expression of this 3.6 kb segment contained in pASP-9.8 but missing from pASP-6.2. In 293T cells, the effect of this 880 bp segment is striking and largely independent of the particular allele at NACP–Rep1. However, in SH-SY5Y cells different alleles had a significant effect on expression levels over a nearly 3-fold range. However, this effect was not simply linear with allele length but appeared to be biphasic in nature. Interestingly, the fact that the effect of the different alleles on the expression level is striking only in the neuroblastoma cells, implies that neuro-specific trans-acting factor/s might be involved in the modulation of expression driven by the NACP–Rep1 element.

When we studied the 880 bp segment in more detail in order to dissect out the contributions to gene expression of various portions of this segment, the two segments that flank the NACP–Rep1 appeared responsible, in an additive manner, for expression while the NACP–Rep1 itself was shown to have a negative effect on expression in both cell lines. These results suggest a model of two regulatory regions, one located upstream of the NACP–Rep1 and the other downstream of the repeat. The transcription factors binding in these two regions may interact to function in an additive manner while the NACP–Rep1 modulates this interaction to a greater or lesser extent depending on which allele is present at the NACP–Rep1 locus. Exactly how the composition of the repeat can modulate this interaction is unknown, but could perhaps be operating by altering how effectively or efficiently the transcription factors that bind upstream and downstream of the repeat can interact. A number of possible models could be envisioned for how the repeat might modulate the interaction between the two flanking regions such as the length of the repeat altering the phase of potential binding sites on the DNA double helix or on chromatin, or the base composition of the repeat altering the flexibility and secondary structure of the DNA, or both.

Five NACP–Rep1 alleles have previously been identified based on a size difference of two nucleotides (8,10) as determined by gel electrophoresis of the PCR products containing the repeats, but the actual base sequence differences responsible for the various alleles was incompletely known. Here, we show that variation in the CA element is most likely responsible for the different length alleles. However, our sequencing analysis of alleles 0 and 1 indicated that the same sequence variant occurred in more than one clone at positions other than within the CA repeat itself. Since such variants were seen more than once, it is possible that the actual number of alleles at this site might be greater than five (Table 1). However, these variants may still simply be the result of the same artifact occurring

repeatedly during the PCR. Distinguishing these possibilities remains a serious challenge and requires comparing large numbers of cloned PCR products from individuals in families and correlating alleles in parents and offspring.

Our study suggests that a microsatellite located a long distance upstream of a gene may play a role in transcriptional regulation of that gene. Several studies in the past have implicated dinucleotide repeats in regulating transcription activity. Expression of the PAX-6 gene (21), the human NRAMP1 gene (22), COL1A2 (23) and MMP-9 (24,25) were shown to be regulated by polymorphic dinucleotide repeats in their 5′ flanking region, and some alleles of these polymorphic repeat sequences were shown to enhance their expression. It is important to note that in most of the genes studied, the combined dinucleotide repeat was at a distance of ≤1.5 kb upstream of the transcriptional start site. Here, we demonstrated that a complex, mixed dinucleotide repeat can influence gene expression from a much longer distance of ~10 kb, such as might be seen with a long-range enhancer or a locus control region. It has been proposed that alternating sequences of purine–pyrimidine has the potential to adopt a Z-DNA conformation (26,27) which was suggested to be involved in the regulation of gene expression (28–30). Indeed, the repeats implicated in gene regulation described above all contain CA or GT dinucleotides of variable length; the NACP–Rep1 site includes both TA and CA dinucleotides, with polymorphic variation in the CA dinucleotide associated with changes in the function of the *SNCA* promoter/enhancer. Further studies to determine if there is non B-DNA conformation of the NACP–Rep1 repeat *in vivo* are of interest.

Three previous studies have reported an association between certain alleles of the NACP–Rep1 locus and sporadic PD (11–13) whereas another study failed to replicate the finding in a Japanese population (14). There are obvious intricacies and difficulties in the analysis and interpretation of association studies of a complex disease trait such as PD when microsatellite loci are used. One problem is that because of the difficulties inherent in sequencing PCR products containing dinucleotide microsatellite repeats, assigning alleles at the NACP–Rep1 locus has relied on determining the size of the PCR products by gel electrophoresis. However, such size determinations are not straightforward. Even when the same PCR primers are used, various instruments and electrophoresis conditions used in different laboratories may lead to variation in the measurements of fragment size and problems with comparing association studies if two laboratories use the same allele number to refer to different alleles. An additional challenge with a complex, mixed microsatellite like NACP–Rep1 is that size alone may not be adequate for defining alleles if two DNA fragments of the same size differ in their composition, having more of one kind of repeat and fewer of another. Two alleles of the same overall length but with different base compositions may have different functional consequences.

We have shown that the various NACP–Rep1 alleles, cloned into the same reporter construct, have different effects on expression of a reporter gene driven by the *SNCA* promoter/enhancer region. Although one can not exclude that the association between the NACP–Rep1 and sporadic PD is the result of another site in linkage disequilibrium with NACP–Rep1, our experiments lend support to the hypothesis that it is the NACP–Rep1 microsatellite itself that has a biological function in the regulation of *SNCA* gene expression. Thus, one can speculate that polymorphism at this microsatellite might be involved in sporadic PD via its effect on the expression of the *SNCA* gene contributing, among other factors, to the risk for disease.

MATERIALS AND METHODS

Computational analysis: a prediction of the promoter/enhancer region

The analysis of potential transcription factor binding sites was performed on the sequence of the 10.7 kb upstream segment of *SNCA* (GenBank accession no. AF163864) using the MatInspector V2.2 software (BCM Gene finder).

Luciferase reporter constructs

A 10.7 kb DNA fragment upstream of the *SNCA* translation start site was amplified from the human PAC 27M07 (GenBank accession no. AF163864, positions 19 040–29 776) as previously described by Touchman *et al.* (9). The full-length promoter/enhancer plasmid is designated pASP-10.7 (Fig. 1). Seven plasmids containing a series of progressive deletions were constructed by using the following restriction sites (Fig. 1): pASP-9.8, *StuI* in the insert (position 19 889) and the *KpnI* site in the vector; pASP-6.2, *PacI* in the insert (position 23 565) and the *KpnI* site in the vector; pASP-4.1, *SacI* in the insert (position 25 637) and the vector; pASP-3.8, *SpeI* in the insert (position 25 955) and the vector; pASP-3.4, *HindIII* in the insert (position 26 404) and the vector; pASP-1.9, *SalI* in the insert (position 27 882) and the *KpnI* site in the vector. The control plasmid, pAS-1.46, which includes only sequences downstream of the transcriptional start site, was constructed by PCR using the forward primer CCTTCTGCCTTTCCACCCTCGTGAG (positions 28 439–28 463) and the reverse primer CCTTTACACCACACTGGAAAACATAAA (8). Next, we removed each promoter construct by restriction endonuclease digestion at the *MluI–XhoI* sites of pCR-XL-TOPO and cloned each into the *MluI–XhoI* sites of the pGL-3 Basic vector (Promega, Madison, WI) which contains the firefly luciferase coding sequence but lacks eukaryotic promoter or enhancer elements. All the plasmids are designated, as mentioned above, as 'pASP-' followed by the insert size.

The plasmids that contain different alleles at the NACP–Rep1 site were constructed by inserting sequence verified segments of the NACP–Rep1 region derived by PCR amplification of genomic DNA from individuals from the CEPH collection carrying alleles 1, 2 or 3, using the forward primer ASP-F TGAAGTTAACCTCCCCTCAATACC and the reverse primer NACP-R AAGAAGACAGCCATCTGCAAGCC (positions 19 897–19 922). Each of the three PCR products was cloned into the pCR-XL-TOPO vector. The sequence of each of the alleles at the NACP–Rep1 site was determined. Next, each construct was cut at the *StuI* site in the insert and the *MluI* site in the vector and cloned into these sites of the pASP-10.7 plasmid, replacing its original *StuI–MluI* fragment. The plasmids carrying the NACP–Rep1 alleles are designated pASP-1, pASP-2 and pASP-3 according to the allele number each carries.

The plasmid harboring the NACP–Rep1 deletion, pASP-ΔNACP (Fig. 1), was constructed using the SOE PCR method (31). The region upstream of the NACP–Rep1 site (positions 19 040–19 355)

was amplified from pASP-1 using the forward primer ASP-F and the reverse primer CTGATGCCTTCCATAGCTACTAATCCATCC which includes a tail of 10 bp complementary to the 5′ of the region downstream the NACP–Rep1. The region downstream of the NACP–Rep1 site (positions 19 439–19 922) was amplified from pASP-1 using the reverse primer NACP-R and the forward primer GTAGCTATGGAAGGCATCAGATATCTCATG, which includes a tail of 10 bp complementary to the 3′ of the region upstream the NACP–Rep1. These two PCR products share a complementary 20 bp fragment. Next, PCR was performed to combine the PCR products of the regions upstream and downstream of the NACP–Rep1 using the forward primer and the reverse primer. The PCR products of the regions upstream and downstream of the NACP–Rep1 were mixed and denatured at 100°C for 10 min and transferred immediately to ice, then added to the PCR mixture and subjected to 30 cycles of PCR as follows: 94°C for 2 min, 55°C for 2 min, 72°C for 2 min, followed by 72°C for 10 min. The combined PCR product was cloned into the pCR-XL-TOPO vector and its sequence and orientation in the vector determined. Next the construct was cut at the *StuI* site in the insert and the *MluI* site in the vector and cloned into these sites of the pASP-10.7 plasmid, replacing its original *StuI–MluI* fragment.

Plasmids carrying a deletion of the NACP–Rep1 and the region upstream of the repeat, pASP-3′NACP, or a deletion of the NACP–Rep1 and a 452 bp fragment downstream of the repeat, pASP-5′NACP (Fig. 1), were constructed as follows. For pASP-3′NACP, the region downstream of the NACP–Rep1 (position 19 439–19 922) was amplified using the reverse primer NACP-R and the forward primer GAAGGCATCAGATATCTCATG. For pASP-5′NACP, the region upstream of the NACP–Rep1 (position 19 040–19 355) was amplified using the forward primer NACP-F and the reverse primer AGGCCTCATAGCTACTAATCCATCC in which a *StuI* site was inserted. Each PCR product was cloned into the pCR-XL-TOPO vector and its sequence was determined. Next, each construct was cut at the *StuI* site in the insert and the *MluI* site in the vector and cloned into these sites of the pASP-10.7 plasmid, replacing its original *StuI–MluI* fragment.

Cell culture and transfection

293T, a transformed human kidney cell, and SH-SY5Y, a human neuroblastoma cell line (American Type Culture Collection), were grown in Dulbecco's modified Eagle's medium (DMEM) (glucose at 4.5 g/l) and DMEM/F-12 1:1 medium, respectively, supplemented with 10% fetal bovine serum, 2 mM glutamine, 100 U/ml penicillin and 100 mg/ml streptomycin. Cells were maintained at 37°C in a humidified 5% CO_2 incubator. A total of 2.5×10^5 293T cells and 1×10^6 SH-SY5Y cells were plated onto each well of a six-well dish the day prior to transfection. To test the expression of each construct, in 293T cells, 100 ng of pASP-10.7, pASP-1, pASP-2 and pASP-3, an amount of each deleted pASP plasmid that was calculated to be a molar equivalent of 100 ng of pASP-10.7, or 33 ng of the pGL 3-Basic plasmid was used. Each test plasmid was mixed with 1 ng of the reference plasmid, pRL-TK (harboring the HSV thymidine kinase promoter upstream of Renilla luciferase), and transfected by the calcium phosphate method using a mammalian transfection kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions. Cells were incubated for 24 h at 37°C, washed with phosphate-buffered saline, and incubated in fresh medium for an additional 24 h. For each cotransfection experiment into SH-SY5Y cells, 2 μg of pASP-10.7, pASP-1, pASP-2 and pASP-3 or the molar equivalent of each of the deleted pASPs plasmids, or 660 ng of pGL 3-Basic and 10 ng of the reference plasmid, pRL-SV40 (harboring the SV40 early enhancer/promoter region upstream of Renilla luciferase), were mixed and cotransfected using the FuGENE™ 6 Transfection Reagent (Roche, Indianapolis, IN) according to the manufacturer's instructions. Cells were incubated for 48 h at 37°C prior to harvesting.

The three wells were independently transfected in parallel with three individually prepared aliquots of transfection reaction and the results from all three replicates were averaged. Each triplicate experiment was repeated three to six times on separate days.

For each construct, three to six experiments were performed. For each construct in each cell line, one experiment consisted of performing the transfection and expression assay in triplicate on three wells of cultured cells independently transfected in parallel with three individually prepared aliquots of transfection reaction.

Luciferase assay

293T and SH-SY5Y cells were washed and lysed in 150 and 200 μl of passive lysis buffer (Promega), respectively. Firefly luciferase and Renilla luciferase activities were measured with 5 μl of 293T or 10 and 20 μl of SH-SY5Y cell lysate using the Dual-Luciferase Reporter assay system (Promega) in a luminometer (EG&G Wallac, Germany). 'Relative activity' was defined as the ratio of firefly luciferase activity to Renilla luciferase activity and was calculated by dividing luminescence intensity obtained in the assay for firefly luciferase by that obtained for Renilla luciferase. 'Fold expression' is defined as the ratio of the relative activity seen with each test plasmid carrying a portion of the *SNCA* upstream region to the basal relative activity and was calculated by dividing the average value of relative activity of each construct to the relative activity of the pGL 3-Basic plasmid without any insert. Overall statistical significance of differences in expression among all the different deletion constructs were analyzed by ANOVA while pairwise comparisons of between different constructs were made with the two-tailed Student's *t*-test (Smith's Statistical Package freeware, http://www.economics.pomona.edu/StatSite/SSP.html).

Genotyping the NACP–Rep1 alleles and sequence determination of the different alleles

For NACP–Rep1 alleles were genotyping by size by previously published methods (10). The region of the NACP–Rep1 polymorphism was amplified by PCR with the primers Rep1 (CCTGGCATATTTGATTGCAA) and Rep2 (GACTGGCCCAAGATTAACCA). The forward primer, Rep1, contained the conjugated fluorophore FAM (Gibco BRL, Gaithersburg, MD). Two microliters of each PCR product, 2.5 μl of formamide, 0.5 μl of loading buffer and 0.5 μl of size standard (Prism Genescan-350 Tamra; Applied Biosystems, Foster City, CA) were denatured at 95°C for 5 min and separated on an ABI 377 automated DNA sequencer with Genescan and Genotyper software. The convention established by Xia *et al.* (8) for naming

alleles was used (PCR product length of 265 bp = allele –1, 267 bp = allele 0, 269 bp = allele 1, 271 bp = allele 2, 273 bp = allele 3).

The determination of the sequences of the different alleles and the basis of the size difference was performed as follows: PCR products from individuals homozygous for alleles 0, 1, 2 (four, six and one individuals, respectively) and an individual compound heterozygous for alleles 1 and 3 were cloned into the pCR-TOPO vector (Invitrogen, Carlsbad, CA) and subsequently sequenced by standard dideoxy chain termination methods using fluorescently labeled nucleotides and the sequence determined on an ABI 377 sequencer (Seqwright).

ACKNOWLEDGEMENTS

We thank Dr Brian Potterf for his help in setting up luciferase assays and Dr Jennifer Johnston for her help with the Genescan analysis. This work was performed within the Intramural Research Program of the National Human Genome Research Institute.

REFERENCES

1. Ueda, K., Fukushima, H., Masliah, E., Xia, Y., Iwai, A., Yoshimoto, M., Otero, D.A., Kondo, J., Ihara, Y. and Saitoh, T. (1993) Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc. Natl Acad. Sci. USA*, **90**, 11282–11286.
2. Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R. *et al.* (1997) Mutation in the α-synuclein gene identified in families with Parkinson's disease. *Science*, **276**, 2045–2047.
3. Kruger, R., Kuhn, W., Muller, T., Woitalla, D., Graeber, S., Kosel, S., Przuntek, H., Epplen, J.T., Schols, L. and Riess, O. (1998) Ala30Pro mutation in the gene encoding α-synuclein in Parkinson's disease. *Nat. Genet.*, **18**, 106–108.
4. Spillantini, M.G., Schmidt, M.L., Lee, V.M., Trojanowski, J.Q., Jakes, R. and Goedert, M. (1997) α-Synuclein in Lewy bodies. *Nature*, **388**, 839–840.
5. Masliah, E., Rockenstein, E., Veinbergs, I., Mallory, M., Hashimoto, M., Takeda, A., Sagara, Y., Sisk, A. and Mucke, L. (2000) Dopaminergic loss and inclusion body formation in α-synuclein mice: implications for neurodegenerative disorders. *Science*, **287**, 1265–1269.
6. Feany, M.B. and Bender, W.W. (2000) A *Drosophila* model of Parkinson's disease. *Nature*, **404**, 394–398.
7. Golbe, L.I. (1999) α-Synuclein and Parkinson's disease. *Mov. Disord.*, **14**, 6–9.
8. Xia, Y., Rohan de Silva, H.A., Rosi, B.L., Yamaoka, L.H., Rimmler, J.B., Pericak-Vance, M.A., Roses, A.D., Chen, X., Masliah, E., DeTeresa, R. *et al.* (1996) Genetic studies in Alzheimer's disease with an NACP/α-synuclein polymorphism. *Ann. Neurol.*, **40**, 207–215.
9. Touchman, J.W., Dehejia, A., Chiba-Falek, O., Cabin, D.E., Schwartz, J.R., Orrison, B.M., Polymeropoulos, M.H. and Nussbaum, R.L. (2001) Human and mouse α-synuclein genes: comparative genomic sequence analysis and identification of a novel gene regulatory element. *Genome Res.*, **11**, 78–86.
10. Hellman, N.E., Grant, E.A. and Goate, A.M. (1998) Failure to replicate a protective effect of allele 2 of NACP/α-synuclein polymorphism in Alzheimer's disease: an association study. *Ann. Neurol.*, **44**, 278–281.
11. Kruger, R., Vieira-Saecker, A.M., Kuhn, W., Berg, D., Muller, T., Kuhnl, N., Fuchs, G.A., Storch, A., Hungs, M., Woitalla, D. *et al.* (1999) Increased susceptibility to sporadic Parkinson's disease by a certain combined α-synuclein/apolipoprotein E genotype. *Ann. Neurol.*, **45**, 611–617.
12. Tan, E.K., Matsuura, T., Nagamitsu, S., Khajavi, M., Jankovic, J. and Ashizawa, T. (2000) Polymorphism of NACP–Rep1 in Parkinson's disease: an etiologic link with essential tremor? *Neurology*, **54**, 1195–1198.
13. Farrer, M., Maraganore, D.M., Lockhart, P., Singleton, A., Lesnick, T.G., de Andrade, M., West, A., de Silva, R., Hardy, J. and Hernandez, D. (2001) α-Synuclein gene haplotypes are associated with Parkinson's disease. *Hum. Mol. Genet.*, **10**, 1847–1851.
14. Izumi, Y., Morino, H., Oda, M., Maruyama, H., Udaka, F., Kameyama, M., Nakamura, S. and Kawakami, H. (2001) Genetic studies in Parkinson's disease with an α-synuclein/NACP gene polymorphism in Japan. *Neurosci. Lett.*, **300**, 125–127.
15. Mezey, E., Dehejia, A.M., Harta, G., Suchy, S.F., Nussbaum, R.L., Brownstein, M.J. and Polymeropoulos, M.H. (1998) α-Synuclein is present in Lewy bodies in sporadic Parkinson's disease. *Mol. Psychiatry*, **3**, 493–499.
16. Sadot, E., Heicklen-Klein, A., Barg, J., Lazarovici, P. and Ginzburg, I. (1996) Identification of a tau promoter region mediating tissue-specific-regulated expression in PC12 cells. *J. Mol. Biol.*, **256**, 805–812.
17. Heicklen-Klein, A. and Ginzburg, I. (2000) Tau promoter confers neuronal specificity and binds Sp1 and AP-2. *J. Neurochem.*, **75**, 1408–1418.
18. Campos-Caro, A., Carrasco-Serrano, C., Valor, L.M., Viniegra, S., Ballesta, J.J. and Criado, M. (1999) Multiple functional Sp1 domains in the minimal promoter region of the neuronal nicotinic receptor α5 subunit gene. *J. Biol. Chem.*, **274**, 4693–4701.
19. Bai, G. and Kusiak, J.W. (1995) Functional analysis of the proximal 5′-flanking region of the *N*-methyl-D-aspartate receptor subunit gene, NMDAR1. *J. Biol. Chem.*, **270**, 7737–7744.
20. West, A., Farrer, M., Petrucelli, L., Cookson, M., Lockhart, P. and Hardy, J. (2001) Identification and characterization of the human parkin gene promoter. *J. Neurochem.*, **78**, 1146–1152.
21. Okladnova, O., Syagailo, Y.V., Tranitz, M., Stober, G., Riederer, P., Mossner, R. and Lesch, K.P. (1998) A promoter-associated polymorphic repeat modulates PAX-6 expression in human brain. *Biochem. Biophys. Res. Commun.*, **248**, 402–405.
22. Searle, S. and Blackwell, J.M. (1999) Evidence for a functional repeat polymorphism in the promoter of the human NRAMP1 gene that correlates with autoimmune versus infectious disease susceptibility. *J. Med. Genet.*, **36**, 295–299.
23. Akai, J., Kimura, A. and Hata, R.I. (1999) Transcriptional regulation of the human type I collagen α2 (COL1A2) gene by the combination of two dinucleotide repeats. *Gene*, **239**, 65–73.
24. Peters, D.G., Kassam, A., St Jean, P.L., Yonas, H. and Ferrell, R.E. (1999) Functional polymorphism in the matrix metalloproteinase-9 promoter as a potential risk factor for intracranial aneurysm. *Stroke*, **30**, 2612–2616.
25. Shimajiri, S., Arima, N., Tanimoto, A., Murata, Y., Hamada, T., Wang, K.Y. and Sasaguri, Y. (1999) Shortened microsatellite d(CA)21 sequence down-regulates promoter activity of matrix metalloproteinase 9 gene. *FEBS Lett.*, **455**, 70–74.
26. Hamada, H. and Kakunaga, T. (1982) Potential Z-DNA forming sequences are highly dispersed in the human genome. *Nature*, **298**, 396–398.
27. Nordheim, A. and Rich, A. (1983) The sequence (dC-dA)n X (dG-dT)n forms left-handed Z-DNA in negatively supercoiled plasmids. *Proc. Natl Acad. Sci. USA*, **80**, 1821–1825.
28. Hamada, H., Seidman, M., Howard, B.H. and Gorman, C.M. (1984) Enhanced gene expression by the poly(dT-dG).poly(dC-dA) sequence. *Mol. Cell. Biol.*, **4**, 2622–2630.
29. Schroth, G.P., Chou, P.J. and Ho, P.S. (1992) Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes. *J. Biol. Chem.*, **267**, 11846–11855.
30. Rich, A. (1994) Speculation on the biological roles of left-handed Z-DNA. *Ann. N. Y. Acad. Sci.*, **726**, 1–17.
31. Warrens, A.N., Jones, M.D. and Lechler, R.I. (1997) Splicing by overlap extension by PCR using asymmetric amplification: an improved technique for the generation of hybrid proteins of immunological interest. *Gene*, **186**, 29–35.

Biochemical and Biophysical Research Communications 286, 381–387 (2001)
doi:10.1006/bbrc.2001.5391, available online at http://www.idealibrary.com on IDEAL®

Identification and Characterization of the Promoter Region of the GRM3 Gene

Corrado Corti,*,†,[1] John H. Xuereb,* Mauro Corsi,† and Francesco Ferraguti*

*Cambridge Brain Bank Laboratory, Department of Pathology, University of Cambridge, Level 3 Laboratory Block Addenbrooke's Hospital, Box 231, Hills Road, Cambridge CB2 2QQ, United Kingdom; and †Department of Biology, GlaxoSmithKline S.p.A. Medicine Research Centre, via Fleming 4, 37135 Verona, Italy

Received July 10, 2001

We have recently described the genomic organisation of the human metabotropic glutamate receptor 3 (GRM3) gene. The putative promoter region is characterised by the presence of a CCAAT and Sp1 site and the absence of a TATA box. Using a reporter gene assay, now we describe the functional activity of GRM3 promoter by transient transfection in both human neuroblastoma and astroglioma cell lines. Deletion of the CCAAT box and Sp1 site resulted in a pronounced reduction of reporter gene expression in both cell types, which indicates that these elements to correspond to the core promoter region. Moreover, we found that the genomic sequence 140 bp upstream of the first transcription initiation site appears to contain regulatory promoter elements for a preferential transcription of the gene in neuroblastoma cells. We also provide evidence that the genomic sequence spanning exon I, corresponding to the GRM3 5′-untranslated region, contains a negative regulatory element that represses gene transcription. © 2001 Academic Press

Key Words: **human; metabotropic; glutamate; receptor; gene; promoter; neuroblastoma; astroglioma.**

In situ hybridisation and immunohistochemical studies in the rat brain have demonstrated that the metabotropic glutamate receptor 3 (mGluR3) is widely, although discretely, distributed in the central nervous system (1, 2). Under physiological conditions mGluR3, unlike many other mGluR, is expressed in both neuronal and glial cells (1–4).

Up-regulation of mGluR3 has been recently reported in reactive astrocytes, both after kainate-induced neuronal injury (5) and in the hippocampus of rats exhibiting behavioural status epilepticus (6). In monoarthritic rats mGluR3 mRNA was also observed to be up-regulated in both neocortex and thalamus (7, 8). Furthermore, the exposure of cultured astrocytes to growth factors or cytokines was shown to induce a marked enhancement of mGluR3 expression (9). These studies suggest a lack of neural specificity for the mGluR3 gene and post-lesional regulation of transcription. However, the genomic mechanisms that regulate transcriptional activity and tissue and/or cell specificity of the mGluR3 gene remain to be identified. As a first step toward identifying these genomic regulatory elements, we have mapped the structure of the human mGluR3 gene (GRM3), which appears to span 220.9 kb in length and to consist of six exons and five introns (10). It is also characterised by the presence of two distinct transcription initiation sites (TISs) on exon I, which are separated by 60 bp (10). A putative promoter region was found, by *in silico* analysis, 260 bp upstream of the first TIS. In common with several other G-protein-coupled receptors (GPCRs), the GRM3 gene lacks a TATA box and CpG islands, but has a CCAAT box and a Sp1 binding site in the vicinity of the TISs (10).

In the present paper, we report that we have cloned and characterised the putative promoter region of the GRM3 gene and examined its transcriptional activity in human astroglioma and neuroblastoma cells using a luciferase reporter gene assay. The minimal sequence maintaining promoter activity was then identified by progressive deletions. In addition, we provide evidence for a strict transcriptional regulation by the 5′-untranslated region (5′-UTR) of the gene.

Abbreviations used: mGluR3, metabotropic glutamate receptor 3; GRM3, human metabotropic glutamate receptor 3 gene; TIS, transcription initiation site; GPCRs, G-protein-coupled receptors; 5′-UTR, 5′ untranslated region; PAC, P1-derived artificial chromosome; PCR, polymerase chain reaction; CHO, Chinese hamster ovary cell; RT-PCR, reverse-transcribed polymerase chain reaction; HD, Huntington's disease.

[1] To whom correspondence should be addressed at Department of Biology, GlaxoSmithKline S.p.A. Medicine Research Centre, via Fleming 4, 37135 Verona, Italy. Fax: +39 0459218 072. E-mail: cc18322@gsk.com.

0006-291X/01 $35.00

MATERIAL AND METHODS

Screening of the human genomic P1-derived artificial chromosome (PAC) library. All oligonucleotide primers used in this work were provided by the Human Molecular Genetic Group, Department of Pathology, University of Cambridge. The human genomic DNA PAC library RPCI1 (kindly provided by Peter de Jong and the Human Genomic Mapping Project, Hinxton) was screened by polymerase chain reaction (PCR) using a set of oligonucleotide primers laying on exon I of the GRM3 gene (For, 5'-CGCTCTCCTAATCTCCCTCTGG-3' and Rev, 5'-CTCCTCTTCTCTTATCAGG-3'). PCR reactions were performed as previously described (11), with minor modifications (initial denaturation at 94°C for 8 min, 0.2 μl *Taq* polymerase, Promega). Control PCR reactions were carried out on human total genomic DNA.

Reporter gene constructs. Poly(A)$^+$ mRNA was extracted from human caudate nucleus (provided by the Cambridge Brain Bank Laboratory, University of Cambridge) by means of a Oligotex suspension (Qiagen). Reverse transcription (RT) reactions were performed with 1 μg of poly(A)$^+$ mRNA using the First-Strand cDNA Synthesis Kit (Pharmacia Biotech). Reporter gene constructs were prepared either by PCR amplification, using the isolated PAC clone 233H20 or human caudate cDNA as templates, or by enzymatic deletion of PCR amplified products. PCR reactions were carried out as described above using 1 μl of *Pfu* polymerase (Promega). Fragment −981/−1 was amplified by PCR from PAC clone 233H20 (primers: For, 5'-TTGCTTGTACGTGGTCAATG-3' and Rev, 5'-GTACGCAGCTCAGCCGTA-3'); fragment −981/+1063 was obtained from two overlapping fragments, one amplified from PAC clone 233H20 (primers: For, 5'-TTGCTTGTACGTGGTCAATG-3' and Rev, 5'-TCTGGCTCATGGTTGGTCCT-3') and one from human caudate cDNA (primers: For, 5'-CTGGCTCACTGAAGACTCT-3' and Rev, 5'-TCTGGCTCATGGTTGGTCCT-3'); fragment −981/−142 was produced by enzymatic deletion (*Eco*RI) of the −981/−1 plasmid; fragment −981/−533 was obtained by deletion of the −981/−1 plasmid with *Xba*I. A construct (−981/+853) containing the promoter region and exon I only was also made by restriction of plasmid −981/+1063 with *Sma*I/*Xho*I. All fragments were then subcloned into the *Nhe*I/*Xho*I sites of the luciferase reporter vector pGL3-Basic (Promega). All amplified PCR products were subjected to nucleotide sequence analysis with a dye terminator cycle sequencing ready-reaction kit (ABI Prism, Perkin Elmer).

Detection of h-mGluR3 in cell lines by RT-PCR. The human astroglioma cell lines, CCRF-C, H4, T98G, Tp265, Tp276, Tp 336, Tp365, Tp410, Tp483, Tp838, U87-MG, U118-MG, U178-MG, U251-MG and U563-MG (kindly provided by Professor V. P. Collins; Department of Pathology, University of Cambridge), the human neuroblastoma cell line SKN-MC, and Chinese hamster ovary cells (CHO) were used to extract poly(A)$^+$ mRNA. Reverse transcriptase and PCR reactions were then performed using the set of oligonucleotide primers: For, 5'-CGCTCTCCTAATCTCCCTCTGG-3' and Rev, 5'-CTCCTCTTCTCTTATCAGG-3'. Poly(A)$^+$ mRNA from human caudate nucleus was used as positive control.

Transient transfections and luciferase assay. For transfection of the h-mGluR3 reporter gene constructs, 3 cell lines were used: COS-7, U178-MG and SKN-MC. COS-7 cells were cultured in high glucose DMEM medium (Gibco); U178-MG astroglioma cells were cultured in F-12 HAM medium (Gibco); and SKN-MC cells were cultured in MEM medium with the addition of nonessential amino acids (Gibco). All media were supplemented with 10% dialysed foetal bovine serum (dFBS, Gibco), 2 mM glutamine, 100 units/ml penicillin and 100 units/ml streptomycin (Gibco). Cells were maintained in a humidified 5% CO_2 atmosphere at 37°C.

Reporter gene constructs were transfected into cell lines by means of Lipofectamine (Gibco) according to the manufacturer's instructions. Briefly, 1.5×10^5 COS-7 and U178-MG or 3×10^5 SKN-MC cells were plated in 6-well tissue culture dishes and transfected the following day, using 0.8 μg of DNA and 4 μl of Lipofectamine/well. To normalise for the transfection efficiency, pRL-RSV (Promega) was cotransfected in a 1:50 ratio with each reporter gene construct. Cells were incubated with the DNA–lipofectamine complex for 4 h, washed and maintained in their respective medium for 48 h. The medium was then removed and cells lysed by application of 250 μl of lysis buffer (Promega) followed by mechanical scraping. *Firefly* and *Renilla* luciferase activity were measured using the Dual Luciferase Reporter Assay System (Promega). In each lysate, the activity of the *Firefly* luciferase was normalised for that of the *Renilla* luciferase. Basal luciferase activity was measured in extract of cells transfected with the pGL3-Basic (Promega). Background activity was determined in cells transfected with 0.8 μg of pCEM-4Z (Promega). Luminescence was measured using a Sirius luminometer (Berthold Detection System).

RESULTS AND DISCUSSION

To identify the genomic *cis*-elements responsible for driving transcription and cell-specific expression of h-mGluR3, we have subcloned the region of the gene between −981 and +1063 and assessed its ability to drive expression of a reporter gene. To this end, we have screened the PAC human genomic library RPCI1 by PCR using a set of oligonucleotide primers laying on exon I (Figs. 1A and 1B). Two positive clones, coded 38J23 and 233H20, were isolated (Fig. 1B). A fragment of 981 bp, flanking the 5' of the previously identified TISs (10) was then amplified from clone 233H20 and subcloned into the *Nhe*I–*Xho*I restriction sites of the pGL3-basic plasmid upstream of the luciferase *Firefly* reporter gene (−981/−1). Two additional fragments, containing the 981 bp putative promoter region and either the untranslated sequence contained in exon I (−981/+853) or the entire GRM3 5'-untranslated region (−981/+1063), were prepared and subcloned into the pGL3-basic plasmid. These plasmids were then transiently transfected into the human astroglioma U178-MG and human neuroblastoma SKN-MC cells to test for their putative promoter activity. These cell lines were selected on the basis of their transcriptional permissibility for the GRM3 gene. To this end, we assessed by RT-PCR the expression of mGluR3 mRNA in 16 different human astroglioma cell lines and the neuroblastoma SKN-MC. CHO cell mRNA was used as negative control whereas mRNA extracted from human caudate nucleus was used as positive control (Fig. 1C). All cell lines tested, except CHO, showed h-mGluR3 mRNA expression (Fig. 1C). U178-MG cells were selected because they showed the highest efficacy of mGluR3 PCR amplification among the astroglioma cell lines.

Transient transfection of the plasmid pGL3 −981/−1 in SKN-MC cells resulted in more than 40-fold increase in luciferase activity relative to the basal activity of the pGL3-basic plasmid (Fig. 2B). In U178-MG cells this plasmid drove a 10-fold increase in luciferase activity over basal (Fig. 2B). Transfection of pGL3 −981/−1 into COS-7 cells gave rise to a luciferase activity (mean ± SEM, 10.7 ± 2.09) similar to that measured in

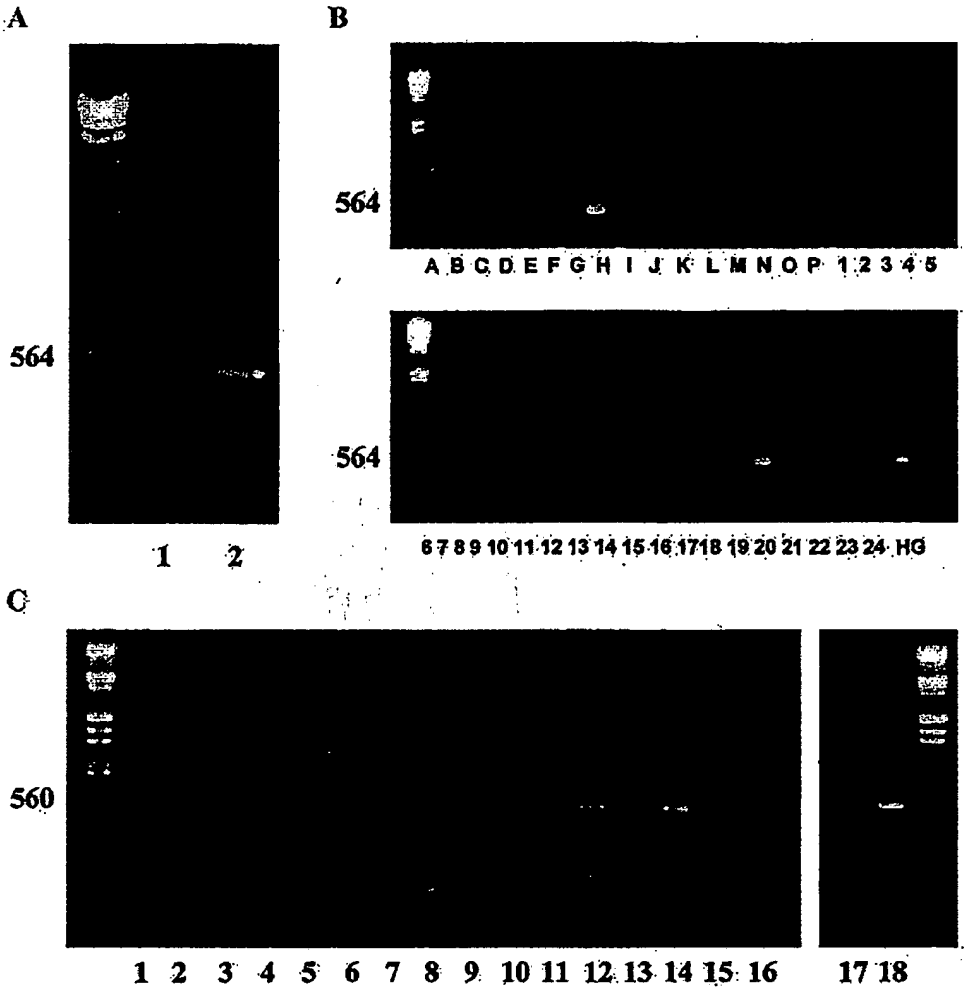


FIG. 1. (A) PCR reaction using oligonucleotide primers laying on exon I of the GRM3 gene on human total genomic DNA. Lane 1, template omitted; lane 2, human genomic DNA. DNA molecular weights are indicated on the left side. (B) Identification by PCR of the PAC clone containing the GRM3 gene. Single clones from a 384-well microtitre plate were pooled in 16 columns and 24 rows. The positive clone in this plate is identified in column H and row 20. Total human genomic DNA was used as positive control (HG). DNA molecular weights are indicated on the left side. (C) RT-PCR analysis of h-mGluR3 expression in different cell lines. Lanes: 1, CCRF-C; 2, CHO; 3, H4; 4, T98G; 5, Tp265; 6, Tp276; 7, Tp 336; 8, Tp 365; 9, Tp 410; 10, Tp483; 11, Tp 838; 12, U87-MG; 13, U118-MG; 14, U178-MG; 15, U251 MG; 16, U563-MG; 17, SKN-MC; 18, human caudate. DNA molecular weights are indicated on the left side.

U178-MG cells. All measures of *Firefly* luciferase activity were corrected for differences in transfection efficiency by co-transfecting with the pRL-SV plasmid, which encodes for *Renilla* luciferase. These results demonstrate that this genomic sequence contains active promoter elements for the expression of the GRM3 gene. The remarkable luciferase activity exhibited by the identified GRM3 promoter in neuroblastoma cells denotes a facilitation of the gene transcription in neuronal cells. However, the relatively abundant reporter gene expression detected in U178-MG astroglioma or COS-7 cells suggests that this genomic region does not confer neuronal specificity to the transcriptional regulation of GRM3. The fact that the expression of mGluR3 mRNA and protein in the rodent and human brain was shown in both glial and neuronal cells (2, 12) is consistent with these results.

To characterise the functional significance of the putative regulatory elements in the GRM3 promoter (10), we deleted 141 bp from the *Eco*RI site to the first TIS (−981/−142) or 532 bp from the *Xba*I site to TIS1 (−981/−533), a region that contains a CCAAT box and a Sp1 binding site consensus sequence (10). Deletion of the 141 bp proximal to TIS1 (−981/−146) produced a marked decrease, approximately 50%, of the transcriptional activity in SKN-MC cells (Fig. 2B). On the other hand, this deletion had no effect on luciferase activity in U178-MG cells (Fig. 2B). Conversely, deletion of the

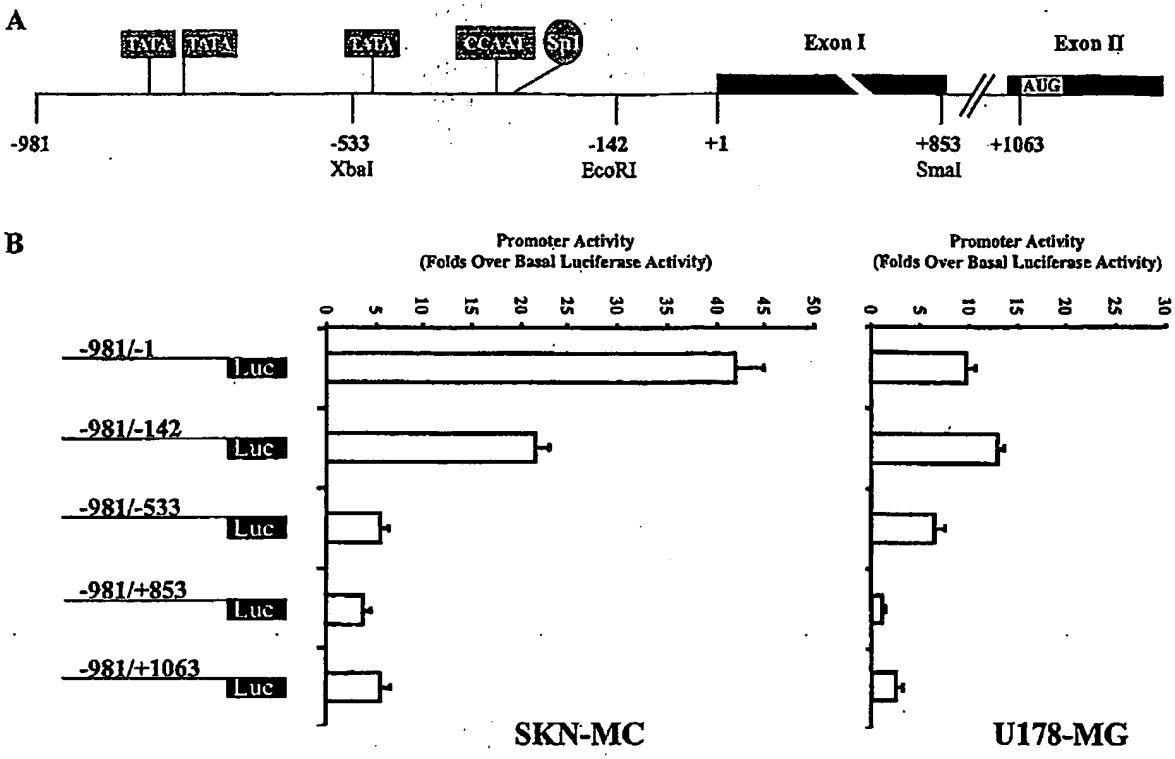


FIG. 2. (A) A schematic representation of the h-mGluR3 promoter region used in reporter gene analysis. Exons are indicated in black boxes and the position of the transcription initiation site is indicated as bp +1 (numeration according to the h-mGluR3 sequence I21435). The TATA boxes, CCAAT box, and Sp1 transcription factor are shown on a grey background. The h-mGluR3 open reading frame is indicated. Restriction sites used to generate deleted constructs are also indicated. (B) Luciferase activity obtained after transfection of the various constructs, indicated on the left side, in SKN-MC human neuroblastoma and U178-MG human astroglioma cells. Results are expressed relatively to the basal luciferase activity obtained with the pGL3-Basic plasmid. Data are normalised for transfection efficiency to *Renilla* luciferase expression driven by cotransfected pRL-RSV. Data represent the mean of all the values obtained from at least three independent experiments performed in triplicate. Data are mean ± SEM.

CCAAT box and Sp1 site (−981/−533), had a very pronounced effect on luciferase activity in both SKN-MC and U178-MG cells, decreasing of approximately 85% and 50% the transcriptional activity of the GRM3 promoter in the two cell lines respectively (Fig. 2B). Thus, the CCAAT box and Sp1 containing region in the GRM3 gene probably corresponds to the core promoter region. It is clear that for many GC-rich promoters, including those of ligand-gated glutamate receptor genes (13), Sp1 elements play a significant role in regulating expression in transfected cells (14). However, it remains to be established whether the Sp1, the CCAAT box or both of these are critical for strong and regulated expression of the GRM3 gene in transfected cells and *in vivo*.

Interestingly, deletion of the 141 bp proximal to the TIS altered the reporter gene transcriptional activity only in neuroblastoma cells. This region is characterised by the presence of a cluster of GATA zinc-finger transcription factor binding sites. GATA transcription factors, which constitute a family composed of several members, direct gene regulation in a variety of discrete cell types (15). It has been proposed that expression of multiple GATA factors in a single cell type may be used to confer specificity in transcriptional activity (15, 16). Hence, it is conceivable that the combination of GATA-1, GATA-3 and GATA-X found in the 5′ proximal region flanking the TIS of the GRM3 gene helps to regulate neuronal-specific expression.

Comparative analysis of the fragments extending into the 5′-UTR of the GRM3 gene showed a dramatic reduction of luciferase activity in both neuroblastoma and astroglioma cells (Fig. 2B). Therefore, the 5′-UTR appears to contain a sequence that exercises transcriptional repression of GRM3. We have shown that the region comprised between +853/+1063 bp does not contribute to the loss of reporter gene expression in the two cell lines tested (Fig. 2B). Consequently, the negative regulatory element is located within exon I. Until now, negative regulation of transcriptional activity by silencers present in the 5′-UTR has been reported only for a small number of genes (17). None of the known position-dependent silencer sequences was present in the GRM3 5′-UTR.

FIG. 3. Nucleotide sequence of the 5′-region comprising the putative promoter and exon I of the GRM3 gene. Nucleotide numbering starts at the first transcription initiation site (+1 according to the h-mGluR3 sequence I21435). Potential regulatory elements, as identified according to the Transcription Factor Binding Sites Databases TRANSFAC and TFSEARCH, are shown underlined and identified by the appropriate symbols. The open arrowheads indicate the position of the two transcription initiation sites. Restriction enzyme sites, indicated by arrowhead, are also shown.

This work is the first to report the identification and functional characterisation of the promoter of a metabotropic glutamate receptor. We have demonstrated that the GRM3 gene is transcribed by a single promoter, which is characterised by the presence of a CCAAT box and an Sp1 binding site in the vicinity of the TIS, and by a silencer situated in the 5′-UTR of exon 1. The GRM3 promoter (Fig. 3) shares many of the characteristic features of G-protein coupled receptor (GPCR) promoters, such as muscarinic (18), dopaminergic (19–21) and 5HT receptors (22, 23). These characteristics include the lack of a TATA box, the presence

of multiple initiation sites and a GC-rich sequence in their proximity. An atypical characteristic observed in the GRM3 promoter region was the presence of 3 TATA-box consensus sequences located approximately 200 and 500 bp upstream of the GC-rich region (Figs. 2A and 3). However, we could not detect any TIS immediately downstream of these TATA boxes (10). Although rather uncommon, this feature has been reported for few other GPCRs (24, 25), but whether they have any functional significance remains to be clarified.

Evidence for a transcriptional regulation of the mGluR3 gene in neuronal and glial cells has been reported recently, both *in vivo* and *in vitro*. In a transgenic mouse model of Huntington's disease (HD), down-regulation of the expression for mGluR3 transcript was described (26). The reduction in mGluR3 mRNA was observed in 4-week-old mice, long before the onset of clinical symptoms, excluding the possibility that this reduction was secondary to cell loss. Thus, the downregulation of mGluR3 expression may represent a primary event in the etiopathogenesis of HD. On the other hand, increased levels of mGluR3 transcript were reported in the cerebral cortex and thalamic reticular nucleus of monoarthritic rats (7, 8). Exposure of cultured astrocytes to basic fibroblast growth factor and epidermal growth factor, which are synthesised and released after brain injury (27), was shown to induce transcription of mGluR3 (9). Consistent with these data, up-regulation of mGluR3 immunoreactivity was observed in reactive astrocytes in the hippocampus after kainate-induced neuronal injury (5) or electrical stimulation of the angular bundle, which induces behavioural status epilepticus (6).

Findings from this study provide the foundation for an understanding of the transcriptional regulation of mGluR3 expression in neurones and glial cells. Future work should seek to confirm a biological role for the identified Sp1, CCAAT and silencer regulatory elements in mediating changes in mGluR3 levels in HD, activated astrocytes and after inflammatory pain.

ACKNOWLEDGMENTS

We thank Dr. Cinzia Sala for helpful comments and criticisms of earlier version of the manuscript. The authors also thank Nicola Hall for her technical help with the luciferase assay and Dr. Carol Sargent for her help with the PAC library screening.

REFERENCES

1. Ohishi, H., Ogawa-Meguro, R., Shigemoto, R., Kaneko, T., Nakanishi, S., and Mizuno, N. (1994) Immunohistochemical localization of metabotropic glutamate receptors, mGluR2 and mGluR3, in rat cerebellar cortex. *Neuron* **13**, 55–66.
2. Petralia, R. S., Wang, Y. X., Niedzielski, A. S., and Wenthold, R. J. (1996) The metabotropic glutamate receptors, mGluR2 and mGluR3, show unique postsynaptic, presynaptic and glial localizations. *Neuroscience* **71**, 949–976.
3. Winder, D. G., and Conn, P. J. (1996) Roles of metabotropic glutamate receptors in glial function and glial–neuronal communication. *J. Neurosci. Res.* **46**, 131–137.
4. Schools, G. P., and Kimelberg, H. K. (1999) mGluR3 and mGluR5 are the predominant metabotropic glutamate receptor mRNAs expressed in hippocampal astrocytes acutely isolated from young rats. *J. Neurosci. Res.* **58**, 533–543.
5. Ferraguti, F., Corti, C., Valerio, E., Mion, S., and Xuereb, J. (2001) Activated astrocytes in areas of kainate-induced neuronal injury upregulate the expression of the metabotropic glutamate receptors 2/3 and 5. *Exp. Brain Res.* **137**, 1–11.
6. Aronica, E., van Vliet, E. A., Mayboroda, O. A., Troost, D., da Silva, F. H., and Gorter, J. A. (2000) Upregulation of metabotropic glutamate receptor subtype mGluR3 and mGluR5 in reactive astrocytes in a rat model of mesial temporal lobe epilepsy. *Eur. J. Neurosci.* **12**, 2333–2344.
7. Lourenco Neto, F., Schadrack, J., Platzer, S., Zieglgansberger, W., Tolle, T. R., and Castro-Lopes, J. M. (2000) Expression of metabotropic glutamate receptors mRNA in the thalamus and brainstem of monoarthritic rats. *Brain Res. Mol. Brain Res.* **81**, 140–154.
8. Neto, F. L., Schadrack, J., Platzer, S., Zieglgansberger, W., Tolle, T. R., and Castro-Lopes, J. M. (2001) Up-regulation of metabotropic glutamate receptor 3 mRNA expression in the cerebral cortex of monoarthritic rats. *J. Neurosci. Res.* **63**, 356–367.
9. Minoshima, T., and Nakanishi, S. (1999) Structural organization of the mouse metabotropic glutamate receptor subtype 3 gene and its regulation by growth factors in cultured cortical astrocytes. *J. Biochem. (Tokyo)* **126**, 889–896.
10. Corti, C., Sala, C. F., Yang, F., Corsi, M., Xuereb, J. H., and Ferraguti, F. (2000) Genomic organization of the human metabotropic glutamate receptor subtype 3. *J. Neurogenet.* **14**, 207–225.
11. Corti, C., Restituito, S., Rimland, J. M., Brabet, I., Corsi, M., Pin, J. P., and Ferraguti, F. (1998) Cloning and characterization of alternative mRNA forms for the rat metabotropic glutamate receptors mGluR7 and mGluR8. *Eur. J. Neurosci.* **10**, 3629–3641.
12. Berthele, A., Platzer, S., Laurie, D. J., Weis, S., Sommer, B., Zieglgansberger, W., Conrad, B., and Tolle, T. R. (1999) Expression of metabotropic glutamate receptor subtype mRNA (mGluR1–8) in human cerebellum. *NeuroReport* **10**, 3861–3867.
13. Myers, S. J., Dingledine, R., and Borges, K. (1999) Genetic regulation of glutamate receptor ion channels. *Annu. Rev. Pharmacol. Toxicol.* **39**, 221–241.
14. Yoo, J., Lee, S. S., Jeong, M. J., Lee, K. I., Kwon, B. M., Kim, S. H., Park, Y. M., and Han, M. Y. (2000) Characterization of the mouse dynamin I gene promoter and identification of sequences that direct expression in neuronal cells. *Biochem. J.* **351**, 661–668.
15. Engel, J. D., Beug, H., LaVail, J. H., Zenke, M. W., Mayo, K., Leonard, M. W., Foley, K. P., Yang, Z., Kornhauser, J. M., Ko, L. J., *et al.* (1992) *cis* and *trans* regulation of tissue-specific transcription. *J. Cell Sci. Suppl.* **16**, 21–31.
16. Lawson, M. A., Buhain, A. R., Jovenal, J. C., and Mellon, P. L. (1998) Multiple factors interacting at the GATA sites of the gonadotropin-releasing hormone neuron-specific enhancer regulate gene expression. *Mol. Endocrinol.* **12**, 364–377.
17. Ogbourne, S., and Antalis, T. M. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem. J.* **331**, 1–14.
18. Buckley, N. J., Bachfischer, U., Canut, M., Mistry, M., Pepitoni, S., Roopra, A., Sharling, L., and Wood, I. C. (1999) Repression and activation of muscarinic receptor genes. *Life Sci.* **64**, 495–499.
19. Minowa, T., Minowa, M. T., and Mouradian, M. M. (1992) Anal-

ysis of the promoter region of the rat D2 dopamine receptor gene. *Biochemistry* **31**, 8389–8396.

20. Minowa, M. T., Minowa, T., Monsma, F. J., Jr., Sibley, D. R., and Mouradian, M. M. (1992) Characterization of the 5′ flanking region of the human D1A dopamine receptor gene. *Proc. Natl. Acad. Sci. USA* **89**, 3045–3049.
21. Beischlag, T. V., Marchese, A., Meador-Woodruff, J. H., Damask, S. P., O'Dowd, B. F., Tyndale, R. F., van Tol, H. H., Seeman, P., and Niznik, H. B. (1995) The human dopamine D5 receptor gene: Cloning and characterization of the 5′-flanking and promoter region. *Biochemistry* **34**, 5960–5970.
22. Parks, C. L., and Shenk, T. (1996) The serotonin 1a receptor gene contains a TATA-less promoter that responds to MAZ and Sp1. *J. Biol. Chem.* **271**, 4417–4430.
23. Zhu, Q. S., Chen, K., and Shih, J. C. (1995) Characterization of the human 5-HT2A receptor gene promoter. *J. Neurosci.* **15**, 4885–4895.
24. Murasawa, S., Matsubara, H., Kijima, K., Maruyama, K., Mori, Y., and Inada, M. (1995) Structure of the rat V1a vasopressin receptor gene and characterization of its promoter region and complete cDNA sequence of the 3′-end. *J. Biol. Chem.* **270**, 20042–20050.
25. Rozen, F., Russo, C., Banville, D., and Zingg, H. H. (1995) Structure, characterization, and expression of the rat oxytocin receptor gene. *Proc. Natl. Acad. Sci. USA* **92**, 200–204.
26. Cha, J. H., Kosinski, C. M., Kerner, J. A., Alsdorf, S. A., Mangiarini, L., Davies, S. W., Penney, J. B., Bates, G. P., and Young, A. B. (1998) Altered brain neurotransmitter receptors in transgenic mice expressing a portion of an abnormal human huntington disease gene. *Proc. Natl. Acad. Sci. USA* **95**, 6480–6485.
27. Schwartz, J. P., Sheng, J. G., Mitsuo, K., Shirabe, S., and Nishiyama, N. (1993) Trophic factor production by reactive astrocytes in injured brain. *Ann. NY Acad. Sci.* **679**, 226–234.

ELSEVIER

Gene 275 (2001) 93–101

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Cloning and characterization of human cathepsin L promoter☆

Radhika Bakhshi[1], Ashish Goel[2], Puneet Seth, Poonam Chhikara, Shyam S. Chauhan*

Department of Biochemistry, All India Institute of Medical Sciences, New Delhi 110029, India

Received 25 January 2001; received in revised form 10 July 2001; accepted 1 August 2001
Received by A.J. van Wijnen

Abstract

Cathepsin L is a lysosomal cysteine protease, which is over-expressed and secreted by malignant cells. It is very potent in degrading collagen, elastin, laminin and other components of the basement membrane and, therefore, has been implicated in tumor invasion and metastasis. The structural portion of the human cathepsin L (*hCATL*) gene was cloned to elucidate its genomic organization (Chauhan et al., J. Biol. Chem. 218 (1993) 1039). In the present study, a 1.90 kb DNA fragment, containing 1825 bp of the 5′ upstream region of *hCATL* and 75 bases of the first exon of the *hCATL*, was amplified by PCR from an adaptor ligated placental genomic library. This fragment has been demonstrated to exhibit promoter activity by luciferase reporter assays. Sequence analysis of this fragment revealed the presence of approximately 29 different putative transcription factor binding sites. Several of them like AP-4, GATA-1, Lmo2, CEBPB, MZF-1, NF-AT, etc. were present more than once in this region. However, a consensus CAAT box but no consensus TATA box was found within the 1.0 kb upstream of exon 1. The transcription initiation site of *hCATL*, using placental total RNA, was mapped to a single adenine residue 289 bases upstream of the ATG codon.

Keywords: TATA-less promoter; Transcription initiation site; Polymerase chain reaction; 5′ rapid amplification of cDNA ends; Reporter gene assay

1. Introduction

Cathepsin L is a potent lysosomal cysteine endopeptidase present in most cell types. In normal cells, its principal function is the degradation and turnover of intracellular proteins. *In vitro*, cathepsin L cleaves a wide range of substrates including extracellular proteins (fibronectin, collagen, elastin and laminin), serum proteins, cytoplasmic proteins and nuclear proteins (Gal and Gottesman, 1986; Mason et al., 1986). It has been shown to play a role in antigen presentation, bone resorption, sperm maturation (Chauhan et al., 1991), and prohormone activation (Marx, 1987). Over-expression of cathepsin L has been reported to be involved in a variety of disease processes such as antigen-induced arthritis, in myofibrillar, myocardial and renal pathologies (Troen et al., 1991), breast tumors and tumor invasion and metastasis (Chauhan et al., 1991).

Cathepsin L is expressed in large quantities by transformed cells in culture (Chauhan et al., 1991). Its synthesis and secretion are also stimulated by tumor promoters (Gottesman and Sobel, 1980), growth factors (Scher et al., 1983), and cAMP (Troen et al., 1991). Over-expression of this protease in non-transformed cells does not confer a transformed phenotype to these cells (Kane et al., 1988), thus suggesting it to be a characteristic of rather than the cause of transformation. Recently, it has also been demonstrated that the carboxy-terminal amino acids of human cathepsin L are essential for its secretion (Chauhan et al., 1998). The expression of cathepsin L was found to be elevated in a variety of human tumors, including cancers of the adrenal gland, breast, colon, kidney, lung, ovary and testicle (Chauhan et al., 1991), in comparison to normal tissues. Increased cathepsin L activity has been linked to advanced tumor stage and grade (Sheahan et al., 1989) and may indicate a poorer prognosis (Lah et al., 1992). Moreover, use of cathepsin L inhibitors has been demonstrated to block invasion by malignant cells *in vitro* (Yagel et al., 1989).

Abbreviations: β-gal, β-galactosidase; bp, base pairs; cAMP, cyclic adenosine 3′,5′-monophosphate; CAT, chloramphenicol acetyl transferase; cDNA, DNA complementary to RNA; DMEM, Dulbecco's modified Eagle's medium; FBS, fetal bovine serum; hCATL, human cathepsin L; hCATL, gene (DNA) encoding CATL; kb, kilobase(s) or 1000 bp; mRNA, messenger RNA; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; TPA, 12-*O*-tetradecanoylphorbol-13-acetate

☆ Sequence data from this article have been deposited with the GenBank Nucleotide Sequence Database under Accession number AF163338.

* Corresponding author. Fax: +91-11-686-2663.
E-mail address: s_s_chauhan@hotmail.com (S.S. Chauhan).
[1] Present address: Department of Pharmacology, Wayne State University, Detroit, MI, USA.
[2] Present address: Johns Hopkins School of Medicine, Baltimore, MD, USA.

0378-1119/01/$ - see front matter
PII: S0378-1119(01)00650-3

Human cathepsin L is encoded by at least four mRNA species, namely hCATL A, AI, AII, and B (Rescheleit et al., 1996). All these mRNA species have identical coding regions; however, differences are found within their 5′ untranslated regions. The hCATL A, AI and AII mRNA are generated by alternative splicing of the same primary transcript. Cloning and determination of the genomic structure of the *hCATL* revealed that hCATL A, its splice variants and hCATL B are transcribed from the same gene which is located on chromosome 9q21-22 (Chauhan et al., 1993). Cloning of the *hCATL* also showed that the 5′ untranslated region of hCATL B is in the 3′ end of the first intron of hCATL A, contiguous to the second exon of this gene.

In mouse fibroblasts cathepsin L mRNA levels are increased by phorbol ester (12-*O*-tetradecanoylphorbol-13-acetate, TPA) and this increase is regulated at the level of transcription (Gottesman and Sobel, 1980; Troen et al., 1991). Similarly in histocytic leukemia cells, it was demonstrated that TPA stimulated cathepsin L gene transcription by a protein kinase independent pathway involving tyrosine kinase activation (Atkins and Troen, 1995). Thus, in view of the elevated hCATL mRNA levels in a variety of human tumors and increased mRNA synthesis in response to TPA and staurosporine it was of interest to study the transcriptional regulation of *hCATL*. Therefore, we cloned human cathepsin L promoter to facilitate the identification of *cis*-regulatory element(s) and *trans*-regulatory factor(s) important for transcriptional regulation of *hCATL* expression.

2. Materials and methods

2.1. Cloning of the 5′ flanking region of hCATL using PCR amplification

The 5′ flanking region of *hCATL* was amplified by PCR using the Promoter Finder™ DNA walking kit (Clontech Laboratories, Inc., Palo Alto, CA). We performed the first long PCR with the Adaptor Primer 1 (AP-1: 5′-CCA TCC TAA TAC GAC TCA CTA TAG GGC-3′, supplied with the kit) and a gene-specific primer SSC 114 (5′-GGC TGG TGC ACA CCT ACT CGA CCG-3′, from exon 1 of *hCATL*, +146 to +123). Subsequently, nested PCR was performed with AP-2 primer (5′-CTA TAG GGC ACG CGT GGT-3′, supplied with the kit) and the nested gene-specific primer SSC 115 (5′-AGT CCC TGT CCG GCC ACA GCT G-3′, +109 to +88). Both primary and nested PCR were performed for 35 cycles on a DNA cycler at 95°C for 30 s, and 68°C for 5 min, with final extension at 68°C for 8 min, after an initial denaturation of 1 min at 95°C. The PCR products were subcloned into the pCR™ II vector (Invitrogen, San Diego, CA) and subjected to double-strand DNA sequencing by the dideoxy chain termination method (Sanger et al., 1977).

2.2. Construction of promoter-luciferase reporter plasmids

A 1.9 kb fragment containing 1825 bp of the 5′ upstream region and the first 75 bp of exon I of *hCATL* was ligated in the *Kpn*I and *Hin*dIII sites of pGL3-Basic vector to obtain the −1825/+75 construct. This construct was transfected in U87MG cells to demonstrate promoter activity in the PCR-amplified 5′ upstream region of *hCATL*.

2.3. Deletion analysis of the promoter

Varying lengths of the 5′ flanking region of *hCATL* were amplified by PCR, using the pTADL1.1 clone as template. A common reverse primer from exon 1 and different forward primers, given below, were used for the amplification of various deletion fragments. *Kpn*I and *Hin*dIII sites were introduced in the forward and reverse primers, respectively, and used for cloning of these fragments upstream of the luciferase reporter gene in pGL3-Basic vector.

For the 5′ ends of the inserts, the following PCR forward primers were used: 5′-ATATGGTACCGCGCCAGGGATCAGAGC-3′ (−133 to −117); 5′-AAGGTACCAGGACTAAGGAATAGCAAGG-3′ (−370 to −352); 5′-TTGGTACCTAGAGAGTGGGCGAATAAAAG-3′ (−783 to −761); 5′-AAGGTACCAAGGTTATGGGGTAAAGGCAGAGG-3′ (−1080 to −1057); 5′-AAGGTACCAGCCTGAACAACAGAGCCA-3′ (−1470 to −1450); 5′-AAGGTACCGGTAAACGTCTTTTCAGG-3′ (−317 to −299); 5′-AAGGTACCTTAGTTGACTCGAGCGCTCC-3′ (−268 to −248); 5′-AAGGTACCACGCCCAGCTCTGGGAC-3′ (−175 to −161); 5′-AAGGTACCTTTAGAAGCAGAGTCAGGCG-3′ (−31 to −12); 5′-AAGGTACCAATCGCAGCGCAGCCAGGCGG-3′ (−86 to −64).

For the 3′ end of the insert, common PCR reverse primer was used: 5′-AATTAAGCTTTCCTGCTGCGGTCGTAGCTGC-3′ (+50 to +75).

The remaining two constructs namely pRB2.75 and pRB4.75 were made by SmaI digestion of the primary construct (pRB1.75) and using erase a base technique (Erase a Base system, Promega, USA), respectively.

2.4. Cell culture

Human glioblastoma cells (U87MG) were obtained from NCCS (National Center for Cell Science, Pune, India) and cultured in DMEM supplemented with 10% FBS and antibiotics under 5% CO_2.

2.5. Cell transfections

U87MG cells were plated in 12-well plates (2×10^5 cells/well) 1 day prior to the transfection. Next day, cells were transfected with TransFast™ reagent (Promega, Madison, WI), as described by the manufacturer. Identical molar amounts of each test or control plasmid (0.7 μg) were mixed with the TransFast™ reagent at a 1:1 ratio in Opti-

MEM reduced medium. pUC19 carrier DNA was used to equalize the amount of the transfected DNA in each sample. DNA-liposome complexes were formed in 0.5 ml of Opti-MEM medium and incubated at room temperature for 15 min. The cells were overlaid with the DNA-liposome complex, after the removal of culture media, and incubated at 37°C in 5% CO_2. Regular growth medium was added after 25 min, containing fetal bovine serum to a final concentration of 10%. Each reporter construct was co-transfected with a 20:1 molar ratio of pCMV-β-gal to normalize for transfection efficiency. Each transfection was performed in quadruplicate.

2.6. Luciferase expression assays

Cells were lysed 48 h post-transfection by two freeze–thaw cycles in 0.1 ml of lysis buffer (Promega) and centrifuged at 10,000 × *g* for 1 min to remove cell debris. The supernatant was used for both luciferase and β-gal activity assays. For luciferase assays, 20 μl of lysate was mixed with 100 μl of luciferase assay reagent (Promega). For the β-gal assay, 10 μl of lysate was mixed with 100 μl of reaction buffer (Clontech). Luciferase and β-gal activities were determined with a MicroLumat LB 96 P luminometer and WinGlow software (EG and G Berthold). Luciferase activity was expressed relative to the β-gal activity and normalized to the value obtained with the promoterless pGL3-Basic vector.

2.7. 5′ rapid amplification of cDNA ends

5′ rapid amplification of cDNA ends (RACE) was performed to identify the transcription start site using a 5′/3′ RACE Kit (Roche GmbH, Mannheim Germany) according to the manufacturer's protocol. First-strand cDNA was synthesized using an antisense gene-specific primer, SSC 160 (5′-CAG TCC CTG TCC GGC CAC-3′) corresponding to +111 to +95 bases in exon 1 of *hCATL* cDNA. The cDNA thus obtained was purified and then terminal transferase was used to add homopolymeric A-tail to the 3′ end of the cDNA. Tailed cDNA was then amplified by PCR using oligo dT anchor primer (supplied in the kit) and another antisense gene-specific primer SSC 165 (5′-GCT TTC CTG CTG CGG TCG TAG CTG GCC AGG-3′) corresponding to +74 to +45 bases in exon 1 of *hCATL* cDNA. The cycling parameters for PCR were: 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min for 30 cycles followed by final extension at 72°C for 10 min. The PCR product (1 μl) was further amplified by a second PCR using the same gene-specific primer, SSC 165, and a PCR anchor primer, supplied in the kit. The cycling parameters were the same as the first PCR. The products were run on 8% polyacrylamide gel and then the 5′ RACE products were purified and cloned in pCR 2.1 vector (Invitrogen, San Diego, CA) and sequenced.

2.8. Statistical analysis

The results were analyzed by one-way analysis of variance followed by the Tukey–Kramer test to identify pairs with statistically significant differences in means.

3. Results

3.1. Cloning of the 5′ flanking region of the human cathepsin L gene

We cloned the *hCATL* promoter by PCR using the Promoter Finder Kit (Clontech). By using this method, two sized fragments of 1.93 and 2.2 kb were amplified from genomic libraries, generated by *Dra*I and *Eco*RV (in different reactions). These fragments were subcloned into a pCR™ II vector to give pTADL1.1 and pTADL3.1, respectively. Upon sequencing with T7 and SP6 primers, pTADL1.1, on one end, exhibited complete homology with the first 112 bases of exon 1 of the *hCATL* gene (Fig. 1A), indicating that the cloned fragment was contiguous to exon 1 of *hCATL*. These results suggested that this clone contained the 5′ flanking genomic region of *hCATL*, which could be the putative promoter. Hence, this clone was used for further studies.

3.2. Promoter activity of the cloned fragment

To determine whether the genomic DNA flanking exon 1 contains a functional promoter, the 1.9 kb fragment of pTADL1.1 containing 1825 bp of the 5′ upstream region of *hCATL* and the first 75 bases of exon 1 was cloned in a luciferase reporter plasmid pGL3-Basic upstream of the luciferase gene. This recombinant plasmid −1825/+75 was transfected into U87MG cells and the cell lysate was assayed for luciferase activity 48 h post-transfection. This construct exhibited more than a 40-fold increase in luciferase activity (see Fig. 3) relative to the promoterless vector (pGL3-Basic). This was also confirmed by CAT assay and CAT ELISA, indicating that the cloned 5′ flanking region of the *hCATL* contains a functional promoter (data not shown).

3.3. Sequence analysis of the promoter region

Both strands of the 1.82 kb promoter region were sequenced using standard DNA sequencing methods and the strategy used is presented in Fig. 1B. The promoter region was shown to have a high GC content. The overall G + C content is 46.9%, while it is 65.9% for the 286 bases proximal to the transcription initiation site (data not shown). Several putative transcriptional elements were identified by the MatInspector v2.2 program (core similarity 0.90, matrix similarity 0.95) using TRANSFAC 4.0 matrices (Quandt et al., 1995; Wingender et al., 2000). Sequence analysis revealed no TATA box within the 1.0 kb region upstream

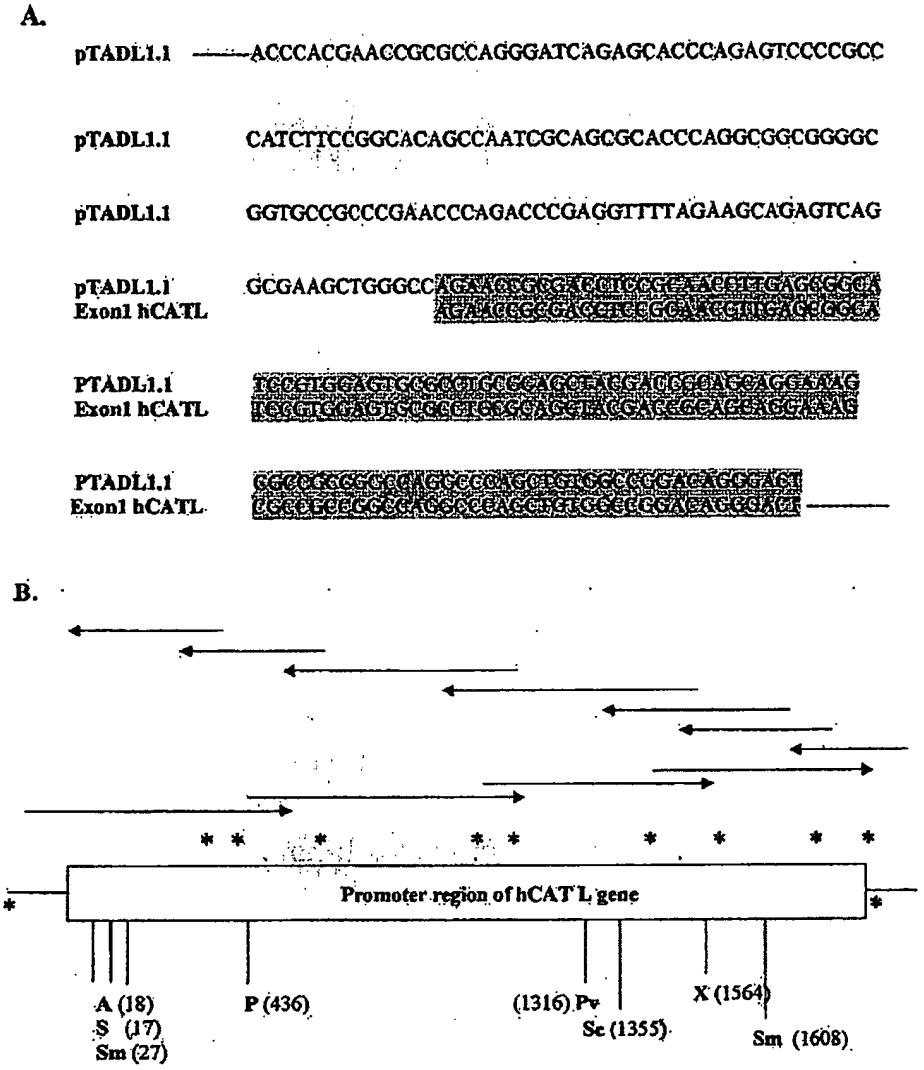


Fig. 1. (A) Sequence homology between clone pTADL1.1 and exon 1 of the human cathepsin L gene. Clone pTADL1.1 containing the 5′ flanking region of *hCATL* was sequenced using T7 and SP6 promoter primers. The partial sequence obtained was aligned with the *hCATL* cDNA, as reported by Gal and Gottesman (1988). The highlighted sequence depicts the homology between the recombinant clone, pTADL1.1, and the 5′ end of exon 1 of the *hCATL*. (B) Sequencing strategy for the promoter region of *hCATL*. Arrows indicate the sequences determined using synthetic oligonucleotides (*) as primers. The figure also represents positions of some common restriction enzymes. A, *AccI*; S, *SalI*; Sm, *SmaI*; P, *PstI*; Pv, *PvuII*; Sc, *SacI*; X, *XhoI*.

of exon 1. Numerous potential binding sites for AP-4, GATA-1, GATA-2, Nkx-2.5, Lmo2 complex, CEBP, etc. and CAAT boxes are present in this promoter region (Fig. 2).

3.4. Identification of the core promoter region

In order to localize the *hCATL* core promoter, and to assess the roles of different *trans*-acting factor binding motifs in the positive or negative regulation of basal *hCATL* expression, sequential unidirectional deletion analysis of the 1.82 kb promoter fragment was carried out. Thus, deletion fragments, having 5′ ends ranging from −1825 to −31 and 3′ ends at +75, were fused with the luciferase gene and transfected into U87MG cells. Luciferase activity was measured in crude cell lysate. As shown in Fig. 3, the plasmid containing the entire 1.82 kb region resulted in a 42-fold higher luciferase activity as compared to the pGL3-Basic vector (i.e. the negative control). Sequential deletion of 355 and 745 bp from the 5′ end of the *hCATL* promoter region resulted in a 2.6-fold ($n = 4$, $P < 0.05$) and 2.3-fold

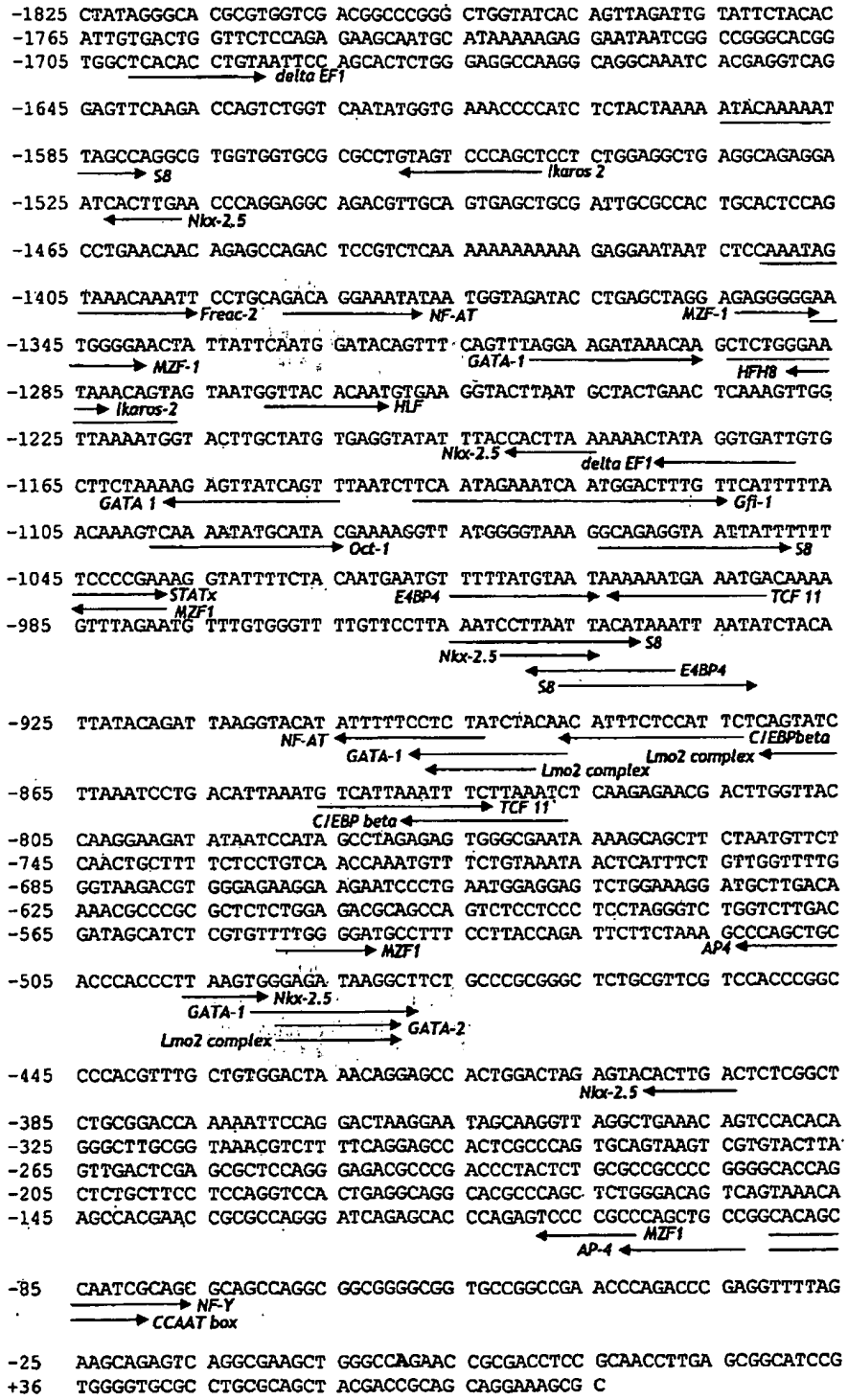


Fig. 2. Nucleotide sequence of human cathepsin L promoter. The nucleotide sequence of the entire 1825 bp of the hCATL promoter region was analyzed by MatInspector v2.2 using the TRANSFAC 4.0 database. The potential regulatory motifs are underlined. The transcription initiation site is shown in boldface.

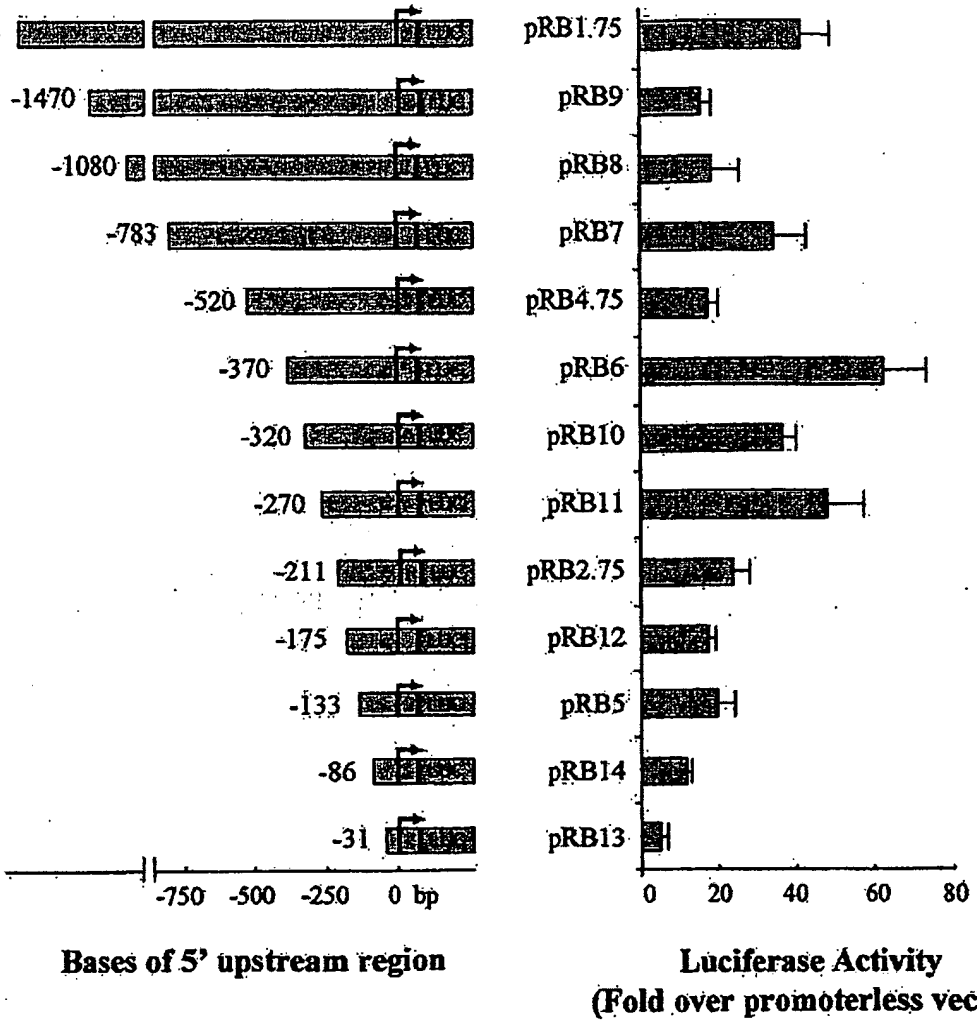


Fig. 3. Deletion analysis of the *hCATL* promoter. The *hCATL* promoter having varying 5′ ends and an identical 3′ end (+75) was fused to the luciferase reporter gene in pGL3-Basic vector as described in Section 2. These constructs were transfected from U87MG cells along with pCMV-β-gal to normalize for transfection efficiency. Luciferase activity is expressed as a fold increase over promoterless reporter plasmid pGL3-Basic (negative control). Results presented are the mean ± SE from four independent experiments.

($n = 4$, $P < 0.05$) decrease in promoter activity, respectively, compared to the construct containing the entire 1.82 kb region (−1825/+75). Upon further deletion of 1045 bases (i.e. −783/+75), a noticeable but statistically non-significant increase with respect to the −1080/+75 construct was observed. Transfection of the −520/+75 fusion construct also exhibited a 2.3-fold decrease ($n = 4$, $P < 0.05$) compared to the −1825/+75 fusion construct. This decrease in reporter gene activity was comparable to the −1470/+75 and the −1080/+75 fusion constructs. These results indicate that motifs present between −1470 and −520 do not play any significant role in the *hCATL* promoter activity in U87MG cells. *HCATL* promoter even after deletion of as many as 1555 bases from the 5′ end (−270/+75 construct) exhibited transcriptional activity comparable to the parent construct (−1825/+75). As little as 133 bases proximal to exon 1 could support transcriptional activity of the reporter gene at 50% levels as compared to the entire 1825 bases of the promoter. Further stepwise removal of sequences spanning between −133 and −31 produced an appreciable drop in the promoter activity. A low promoter activity around five-fold compared to the promoterless pGL3-Basic was observed with the construct containing even 31 bases of the promoter (Fig. 3). Deletion of a cluster of Nkx-2.5, GATA-1, GATA-2 and Lmo2 complex around −495 bp resulted in a statistically significant ($n = 4$, $P < 0.01$) increase in promoter activity when compared to the −520/+75 construct.

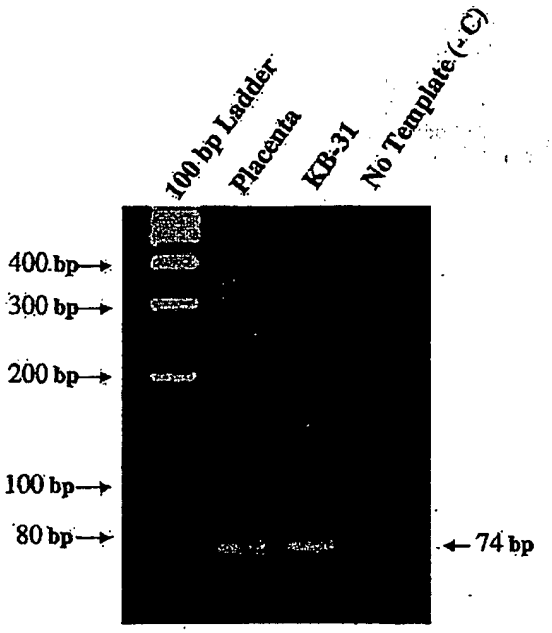


Fig. 4. Mapping of the transcription initiation site of human cathepsin L mRNA by 5′ RACE. Placental total RNA was subjected to 5′ RACE as described in Section 2 and the RACE products were analyzed on 8% acrylamide gel.

3.5. Determination of the transcriptional initiation site

5′ RACE was employed for the mapping of the transcription initiation site of the hCATL mRNA and the 3′ boundary of the *hCATL* promoter. Total RNA from human placenta was reverse-transcribed and tailed at the 3′ end. The tailed cDNA was subjected to PCR using a PCR anchor primer and an antisense gene-specific primer, complementary to the first exon of *hCATL*, from position +74 to +45. This resulted in the amplification of three fragments of sizes 60, 74 and 120 bp approximately (Fig. 4). Six clones of 60, 74 and 120 bp each were subjected to nucleotide sequencing. Sequence analysis of 60 and 120 bp fragments revealed no homology with *hCATL* cDNA reported by Gal and Gottesman (1988). However, the 74 bp fragment exhibited 100% homology with the cDNA of *hCATL* from +1 to +74 bp (data not shown). The *hCATL* cDNA reported by Gal and Gottesman (1988) contain 289 bp upstream of the ATG. These results demonstrate that the transcription initiation site of hCATL mRNA is from 289 bp upstream of the translation start codon.

4. Discussion

Chauhan et al. (1993) cloned the structural portion of the *hCATL* by PCR; however, the promoter region of this gene and its transcriptional regulation remained unknown. In the present study, we have isolated and characterized the human cathepsin L promoter to facilitate the identification of the *cis*-regulatory elements and the *trans*-regulatory factors, which are important for transcriptional regulation of *hCATL* expression. Further, the basic features of this promoter were characterized, including the mapping of the transcription start site and the determination of the core promoter region for the transcriptional activation of the *hCATL*.

The promoter region of *hCATL* was cloned by PCR when repeated attempts to fish out the upstream region from a placental genomic library with a radiolabeled cDNA probe provided false positives (data not shown). PCR of an adaptor ligated human genomic library, using adaptor primers and gene-specific primers, amplified two fragments of sizes 1.93 and 2.2 kb. The cloning and partial sequencing of the 1.93 kb fragment showed the presence of the first 112 bases of the first exon of *hCATL*. These results indicated that this cloned fragment was physically linked to exon 1 of *hCATL* and contained the 5′ flanking region.

The functionality of the putative promoter region of *hCATL* was tested by luciferase assay. Sequence analysis of the 1.0 kb promoter region proximal to the transcriptional start site revealed that it does not contain a consensus TATA box element and has a high G + C content. These features are common to both housekeeping and non-housekeeping gene promoters, such as the promoters of β1-integrin (Cervella et al., 1993) and of Dr-nm23, a member of the putative metastatic suppressor mn23-Hi gene family (Martinez et al., 1997). Other lysosomal enzymes with similar functions, such as cathepsin B (Yan et al., 2000), cathepsin S (Shi et al., 1994) and cathepsin D (Redecker et al., 1991), like cathepsin L also possess TATA-less promoters. These characteristics may be common among the promoter regions of lysosomal enzymes, perhaps indicating that they may be regulated by common mechanisms at the transcriptional level.

TATA-less promoters were originally thought to be constitutive (i.e. expressed at a relatively constant level in all tissues and unregulated); however, the expression of *hCATL* is induced by transformation, tumor promoters and growth factors (Troen et al., 1991). There are now several examples of TATA-less promoters which are differentially regulated during development in a tissue-specific way or in response to viral or pharmacological stimuli (Azizkhan et al., 1993). Though the primary function of cathepsin L is intracellular protein degradation and turnover, it has also been suggested to have specialized roles in certain tissue types, such as antigen presentation, and sperm maturation (Chauhan et al., 1991). Since TATA-less promoters have been shown to be differentially regulated, expression of cathepsin L at a basal level in most cells for its general function of intra-lysosomal protein degradation and at elevated levels in specific tissue types which require higher expression of this enzyme can be explained at the transcriptional level. In this context, abnormal amounts of cathepsin

L found in pathological conditions, such as glomerulonephritis, arthritis and cancer, could result from an ectopic expression of transcription factors that normally regulate the levels of this enzyme during development or in a tissue-specific way. This also seems to be the case for the aspartic protease cathepsin D, which is frequently overexpressed in breast cancer and is induced by estrogens in hormone responsive breast cancer cells (Cavailles et al., 1993).

The transcription initiation site of *hCATL* was determined by 5′ RACE. In human placenta, the hCATL transcript(s) initiated from a single adenine residue located 289 bp upstream of the translation start codon. This site is the same as the previously reported 5′ end of *hCATL* cDNA (Gal and Gottesman, 1988). As GC-rich sequences are often reported in TATA-less promoters (Bird, 1986), the analysis of the promoter region revealed that it is a TATA-less promoter and has a high GC content, similar to that of cathepsins B and D (Yan et al., 2000; Redecker et al., 1991). Potential binding sites for known transcription factors were also identified in the *hCATL* promoter. These included AP-4, CAAT box, GATA-1, GATA-2, MyoD, delta EF1, C/EBP beta, MZF-1, NF-AT, etc. Regulatory elements like AP-1 and AP-2 are normally present in promoters associated with genes which are controlled by phorbol diesters or the growth factor receptor pathway (Nicholson et al., 1990). Although human cathepsin L expression is stimulated by tumor promoters like TPA (Atkins and Troen, 1995), neither AP-1 nor AP-2 elements were found in the entire 1.82 kb region of *hCATL* promoter. Consistent with this observation, treatment of U87MG cells transfected with −1825/+75 luciferase construct with TPA did not exhibit any induction in luciferase activity as compared to the untreated control (data not shown). In a previous study it has been shown that induction of mouse cathepsin L promoter by phorbol esters is mediated by the sequences present in the first intron (Troen et al., 1991). Similar sequences present in the region downstream of the first exon may be responsible for conferring TPA responsiveness to human cathepsin L promoter. We observed up to 60-fold activity with human cathepsin L promoter as compared to promoterless luciferase vector (Fig. 3). This activity was approximately one-third of the luciferase activity obtained with a strong viral promoter SV-40 (data not shown) in U87MG cells. Thus, our results demonstrate high transcriptional activity of *hCATL* promoter in malignantly transformed U87MG cells. Therefore, this cloned promoter can be used to study the molecular basis of elevated expression of *hCATL* in malignancy.

Luciferase assays indicated that the reporter gene construct, containing the entire 1.82 kb promoter region, i.e. pRB1.75, exhibited 42-fold higher luciferase activity than the promoterless pGL3-Basic. Deletions of the promoter region between −1825 and −1470 resulted in a significant decrease in the promoter activity, as compared to pRB1.75. This reflects the presence of putative positive regulatory element(s) in this region. The presence of putative binding sites for delta EF1, S-8, Ikaros-2, and Nkx-2.5 in this region and the subsequent decrease of promoter activity associated with the deletion of this region suggests that this upstream portion of *hCATL* promoter is important for transcriptional activity in U87MG cells. Further stepwise removal of nucleotides from −1470 to −520 did not significantly affect the promoter activity as compared to the −1470/+75 construct (Fig. 3). However, several motifs like Freac-2, NF-AT, MZF-1, GATA-1, HFH-3, E4BP4, Oct-1, Gfi-1, etc. were present in this region. Thus, it can be inferred that either these sequences present between −1470 and −520 are not important for promoter activity or they are only functional when sequences between −1825 and −1470 are present. The proximal 133 bp of the *hCATL* promoter conferred a moderate promoter activity of approximately 50% of the entire 1825 bp region. This region contains AP-4, MZF-1, NF-Y motifs and a CAAT box. Further stepwise removal of 5′ sequences spanning from −133 to −31 produced an appreciable drop in the promoter activity, indicating that this 133 bp fragment proximal to the transcriptional site is a core functional promoter, essential for the basal transcriptional activation of *hCATL*.

The structural part of the *hCATL* was cloned in 1993 (Chauhan et al., 1993). In the present study, we have cloned and sequenced the *hCATL* promoter, defined the core promoter region and determined the transcription initiation site. Sequence analysis and deletion studies indicated the presence of putative binding sites for various transcriptional factors, like AP-4, GATA-1, NF-Y, NF-AT, STATx and CAAT box. Current studies in our laboratory are focusing on the identification of specific elements involved in the transcriptional regulation of *hCATL*.

Acknowledgements

The authors would sincerely like to thank Professor Bonnie F. Sloane (Department of Pharmacology, Wayne State University, Detroit, MI, USA), in whose laboratory a part of this work was carried out (supported therein by US Public Health Service Grant CA 36481 to B.F.S.). We would also like to thank Vinay S. Mahajan and Dr K. Ganeshan for their help in the preparation of this manuscript. We are also thankful to the Council of Scientific and Industrial Research (CSIR), New Delhi for financial assistance in the form of a research grant (S.S.C.) and fellowships (R.B., A.G., P.S. and P.C.).

References

Atkins, K.B., Troen, B.R., 1995. Phorbol ester stimulated cathepsin L expression in U937 cells. Cell Growth Differ. 6, 713–718.

Azizkhan, J.C., Jensen, D.E., Pierce, A.J., Wade, M., 1993. Transcription from TATA less promoters: dihydrofolate reductase as a model. Crit. Rev. Euk. Gene Exp. 3, 229–254.

Bird, A.P., 1986. CpG rich islands and the function of DNA methylation. Nature 321, 209–213.

Cavailles, V., Augereau, P., Rochefort, M., 1993. Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only TATA dependent transcription in breast cancer cells. Proc. Natl. Acad. Sci. USA 90, 203–207.

Cervella, P., Silengo, L., Pastore, C., Altruda, F., 1993. Human β1-integrin gene expression is regulated by two promoter regions. J. Biol. Chem. 268, 5148–5155.

Chauhan, S.S., Goldstein, L.J., Gottesman, M.M., 1991. Expression of cathepsin L in human tumors. Cancer Res. 51, 1478–1481.

Chauhan, S.S., Popescu, N.C., Ray, D., Fleischmann, R., Gottesman, M.M., Troen, B.R., 1993. Cloning, genomic organization and chromosomal localization of human cathepsin L. J. Biol. Chem. 218, 1039–1045.

Chauhan, S.S., Ray, D., Kane, S.E., Willingham, M.C., Gottesman, M.M., 1998. Involvement of carboxy terminal amino acids in the secretion of human lysosomal protease cathepsin L. Biochemistry 37, 8584–8594.

Gal, S., Gottesman, M.M., 1986. The major excreted protein of transformed fibroblasts is an activable acid protease. J. Biol. Chem. 262, 1760–1765.

Gal, S., Gottesman, M.M., 1988. Isolation of a cDNA for human pro-(cathepsin L). Biochem. J. 253, 303–306.

Gottesman, M.M., Sobel, M.E., 1980. Tumor promoters and Kirsten sarcoma virus increase synthesis of a secreted glycoprotein by regulating levels of translatable mRNA. Cell 19, 449–455.

Kane, S.E., Troen, B.R., Gal, S., Ueda, K., Pastan, I., Gottesman, M.M., 1988. Use of a cloned multidrug resistance gene for co-amplification and overproduction of major excreted protein, a transformation-regulated secreted acid protease. Mol. Cell. Biol. 8, 3316–3321.

Lah, T.T., Kokalj-Kunovar, M., Strukelj, B., Pungerčar, J., Barlic-Maganja, D., Drobnic-Kosovok, M., Kastelic, L., Babnik, J., Golouh, R., Turk, V., 1992. Stefins and lysosomal cathepsins B, L and D in human breast carcinoma. Int. J. Cancer 50, 35–44.

Martinez, R., Venturelli, D., Perrotti, D., Veronese, M.L., Kastury, K., Druck, T., Huebner, K., Calabretta, B., 1997. Gene structure, promoter activity and chromosomal location of the DR-nm23 gene, a related member of the nm23 gene family. Cancer Res. 57, 1180–1187.

Marx, J.L., 1987. A new wave of enzymes for cleaving prohormones. Science 235, 285–286.

Mason, R.W., Johnson, D.A., Barrett, A.J., Chapman, H.A., 1986. Elastinolytic activity of human cathepsin L. Biochem. J. 233, 925–927.

Nicholson, R.C., Mader, S., Nagpal, S., Leid, M., Rochette, E.C., Chambon, P., 1990. Negative regulation of the rat stromelysin gene promoter by retinoic acid is mediated by an API binding site. EMBO J. 9, 4443–4454.

Quandt, K., Frech, K., Karas, H., Wingender, E., Werner, T., 1995. MatInd and MatInspector – new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 23, 4878–4884.

Redecker, B., Heckendorf, B., Grosch, H.W., Mersmann, G., Hasilik, A., 1991. Molecular organization of the human cathepsin D gene. DNA Cell Biol. 10, 423–431.

Rescheleit, D.K., Rommerskirch, W.J., Wiederanders, B., 1996. Sequence analysis and distribution of two new human cathepsin L splice variants. FEBS Lett. 394, 345–348.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Scher, C., Dick, P.L., Whipple, A.P., Locatell, K.L., 1983. Identification of a BALB/c-3T3 cell protein modulated by platelet-derived growth factor. Mol. Cell. Biol. 3, 70–81.

Sheahan, K., Shuja, S., Murnane, M.J., 1989. Cysteine protease activities and tumor development in human colorectal carcinoma. Cancer Res. 49, 3809–3814.

Shi, G.P., Webb, A.C., Foster, K.E., Knoll, J.H.M., Lemere, C.A., Munger, J.S., Chapman, H.A., 1994. Human cathepsin S: chromosomal localization gene structure and tissue distribution. J. Biol. Chem. 260, 11530–11536.

Troen, B.R., Chauhan, S.S., Ray, D., Gottesman, M.M., 1991. Downstream sequences mediate induction of the mouse cathepsin L promoter by phorbol esters. Cell Growth Differ. 2, 23–31.

Wingender, E., Chen, X., Hehl, R., Karas, H., Liebich, I., Matys, V., Meinhardt, T., Prüß, M., Reuter, I., Schacherer, F., 2000. TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res. 28, 316–319.

Yagel, S., Warner, A.H., Nellans, H.N., Lala, P.K., Waghorne, C., Denhardt, D.T., 1989. Suppression by cathepsin L inhibitors of the invasion of amnion membranes by murine cancer cells. Cancer Res. 49, 3553–3557.

Yan, S., Berquin, I., Troen, B., Sloane, B., 2000. Transcription of human cathepsin B is mediated by Sp1 and Ets family factors in glioma. DNA Cell Biol. 19, 79–91.

0013-7227/01/$03.00/0
Printed in U.S.A.

Endocrinology 142(9):3987–3995

Cloning and Characterization of the Proximal Murine *Phex* Promoter

SHIGUANG LIU, RONG GUO, AND L. DARRYL QUARLES

Department of Medicine, Duke University Medical Center, Durham, North Carolina 27710

Phex is an endopetidase that regulates systemic phosphate homeostasis. We investigated _Phex_ gene transcription by cloning and performing functional analysis of the 2736 bp of the 5′ flanking region of the mouse _Phex_ gene containing its promoter. We identified a transcription start site, a consensus TATA-box, and multiple potential _cis_-acting regulator elements. To determine whether the promoter directs cell-type restricted _Phex_ expression, we transfected full-length and 5′-deleted _Phex_ luciferase reporter constructs into various cell lines. _Phex_-expressing C5.18 chondrocytes displayed the highest activity of the transfected _Phex_ promoter constructs compared with non-_Phex_-expressing COS-7 cells, whereas promoter activity was intermediate in ROS 17/2.8 osteoblasts and maturation stage-dependent in MC3T3-E1 osteoblasts. Analysis of sequential 5′-deletion mutants of the _Phex_ promoter in ROS 17/2.8 cells revealed bimodal activity, suggesting that both positive and negative _cis_-acting regions may be present. The chondrogenic factor SOX9 markedly stimulated _Phex_ promoter activity, whereas _Cbfa1_, PTH, and $1,25(OH)_2D_3$ had no effect. Our findings are consistent with the predominant expression of _Phex_ in bone and cartilage. Additional studies will be needed to confirm the regulatory regions in the _Phex_ promoter that function in a cell-restricted manner. (_Endocrinology_ 142: 3987–3995, 2001)

PHEX IS A zinc metalloproteinase whose physiologic function is to regulate unidentified substrates controlling circulating phosphate concentrations and skeletal mineralization (1, 2). *Phex* is expressed in the skeleton, as well as in a limited number of other tissues including lung, ovary, testis, parathyroid gland, and brain (3–6). Investigations of the hereditary diseases X-linked hypophosphatemia and the *Hyp* mouse animal homologue, which have inactivating *Phex* mutations, have provided insights into the complex pathogenesis of phosphaturia and impaired skeletal mineralization (1, 2, 7). It is likely that inactivating mutations of *Phex* indirectly causes renal phosphate wasting and hypophosphatemia by its failure to metabolize a putative phosphate-regulating hormone, phosphatonin, that has not yet been identified (2, 8, 9). In addition, *Phex* seems to play a role in regulating skeletal mineralization independent of hypophosphatemia, as evidenced by *Phex* expression in mature osteoblasts (3, 5), the finding that the loss of *Phex* function in osteoblasts derived from *Hyp* mice causes an intrinsic defect in mineralization of extracellular matrix *in vitro* (10, 11) and the partial rescue of the phenotype by bone marrow transplantation (12).

Limited information is available regarding the regulation of *Phex*. Given its role in systemic phosphate homeostasis and skeletal mineralization, and the multitude of factors regulating mineral ion homeostasis, such as age, sex, diet, PTH, and vitamin D, it is likely that *Phex* expression is regulated (2). Recent studies demonstrate that *Phex* expression in bone is increased by IGF-I and down-regulated by PTH administration and aging (13). Other studies indicate that a low phosphorus diet increases *Phex* transcripts in the pituitary gland but not in bone as assessed by semiquantitative RT-PCR analysis (14). In the skeleton, *Phex* is one of a group of cell-restricted genes that displays a late developmental expression pattern in mineralizing bone and cartilage, where this endopeptidase is predominantly expressed in mature osteoblasts and hypertrophic chondrocytes that are involved in regulation mineralization of extracellular matrix (3, 10, 15). Thus, cell-type and/or tissue-specific factors are likely responsible for the restricted pattern of *Phex* expression. Additional studies are needed to characterize molecular mechanisms of cell-type and tissue-specific expression of *Phex* and its regulation by factors controlling mineral homeostasis. Identification and characterization of the *Phex* gene promoter are important steps in defining the role of transcriptional regulation of the *Phex* gene.

In the current studies, we describe the cloning, sequencing, and functional analysis of the murine *Phex* promoter. We found that osteoblasts and chondrocytes express *Phex* endogenously and have the *trans*-acting factors necessary to direct transcription of the transfected *Phex* promoter. In contrast, COS-7 cells lack the necessary transcriptional machinery needed for *Phex* expression, suggesting that regulatory regions of the murine *Phex* promoter may function in a cell-specific manner. SOX9, a master regulatory factor for chondrocyte differentiation, stimulated the promoter activity, whereas the osteoblastic differentiation factor *Cbfa1* had no effect on *Phex* promoter activity. We failed to identify either $1,25(OH)_2D_3$ or PTH regulation of the transfected *Phex* promoter, suggesting any effects of these agents on *Phex* may not be mediated by transcriptional mechanisms or that the *Phex* promoter region lacks the necessary regulatory elements.

Materials and Methods

Isolation of the murine Phex 5′-flanking region

A mouse 129 SVJ genomic library (Stratagene, La Jolla, CA) was screened with a cDNA probe derived from mouse *Phex* exon 1 (oligo-

Abbreviations: G3PDH, Glyceraldehyde-3-phosphate dehydrogenase; RT, reverse transcriptase; TPA, 12-O-tetradecanoylphorbol-13-acetate; UTR, untranslated region.

nucleotides +432 to +547). The probe was labeled with α-^{32}P dCTP (800Ci/mmol) using a random primer DNA labeling kit (Life Technologies, Inc., Grand Island, NY). Plaque lifting, prehybridization, hybridization, washing of the filters, and autoradiography were performed by standard methods. Nine positive clones were identified after screening 1×10^6 pfu at a density of 5×10^4 pfu/plate. Putative positive clones were purified by secondary and tertiary rounds of screening. Restriction digests with *Nsi* I and Southern blot analysis with the exon 1 cDNA probe identified two clones containing identical 5′-flanking regions. A 12-kb fragment was released from this phage clone and partially sequenced, revealing 1104 bp of intron 1, 674 bp exon 1, and approximately 10.2 kb of the 5′-flanking sequence. This promoter sequence has been entered in GenBank, accession number AF299334.

The promoter analysis program Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html) was used to predict the transcription start site. The promoter sequence was analyzed by the Transcription Element Search System (TESS of the Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania, http://agave.humgen.upenn.edu/tess/index.html) by using a 6-bp minimum element size limit, a 5% mismatch allowance, a minimum log-likelihood of homology of 10, and a secondary log-likelihood density threshold of 1.6. Further analysis of the promoter sequence was conduced using TFSEARCH (http://pdap1.trc.rwcp.or.jp/research/db/TFSEARCH.html). In addition, we used the sequence information from the mouse *Phex* promoter and Blast analysis (http://www.ncbi.nlm.nih.gov/BLAST/) to identify the corresponding region of the human promoter (accession number Y10196) (16). The respective regions of the mouse *Phex* and human PHEX promoter were compared using TRES Transcription Regulatory Element Search, a tool for Comparative Promoter Analysis (http://www.bic.nus.edu.sg:8888/tres/) and BestFit analysis.

RT-PCR analysis

RT-PCR was performed using a two-step RNA PCR kit (Perkin-Elmer Corp., Branchburg, NJ). DNase-treated total RNA (2.5 μg) was reverse transcribed into cDNA in a total volume of 50 μl with random primers. The RT reaction was incubated at 42 C for 15 min. The resulting cDNA was PCR amplified using various sets of primers. Forward primers included *Phex*-22F (5′-GGGACTAACACACTGAAAGAGT-3′) and *Phex*+1F (5′-AACTTTTGACGACGACAGTTCA-3′). The reverse primer was *Phex*+697R (5′-GAAACTTAGGAGACCTTGAC-3′), located in exon 2. PCR was performed with thermal cycling parameters of 94 C for 30 sec, 60 C for 30 sec, and 72 C for 60 sec for 35 cycles, followed by a final extension at 72 C for 7 min. For evaluation of *Phex* expression in different cell lines, we used the forward primer *Phex*+1866F (5′-AATTGATTGAGGGTGTTCGC-3′) and the reverse primer *Phex*+2943R (5′-ACCCAAATAATGAAAATGCA-3). For RT-PCR analysis of 4 and 9-d-old MC3T3-E1 osteoblasts, we used *Phex*-specific primers *Phex*+1419F (5′-TTGGCAAAAGTTGGCTATCCAG-3′) and *Phex*+965R (5′-TATCCATTTCCTGTAAGCCC-3′) as described previously (3). Samples without RT treatment were analyzed as controls. For evaluation of SOX9 expression, we used the forward primer 5′-ATCTGAAGAAGGAGGAGCGAG-3′ and the reverse primer 5′-TCAGAAGTCTCCAGAGCTTG-3′, which are to regions conserved across species including mouse and primates (17). Glyceraldehyde-3-phosphate dehydrogenase (G3PDH) or mouse β-actin were amplified as controls for the amounts and integrity of RNA in the PCR reactions. RT-PCR products were identified by autoradiography using radiolabeled *Phex* and β-actin cDNA probes as described previously (3) or by ethidium bromide staining.

RNase protection analysis

RNase protection assays were used to estimate the *Phex* mRNA transcription start site. The riboprobe (−130, +132) consisted of a 262-bp fragment of the mouse *Phex* subcloned into pBSK(−) (Stratagene). Single-stranded antisense radiolabeled RNA probes were transcribed from *Bam*HI linearized *Phex* 262/pBSK(−) using T7 RNA polymerase and [^{32}P-]UTP (NEN Life Science Products, Boston, MA). The riboprobe was purified on a 6% polyacrylamide gel. RNase protection assays were conducted using the RPA III kit (Ambion, Inc., Austin, TX). Twenty micrograms of total RNA from each cell line and the labeled riboprobe (100,000 cpm) were precipitated, redissolved in 10 μl Hybridization III Buffer and incubated overnight at 42 C. Products were digested with RNase A/RNase T1 mix (1:100 dilution) for 30 min, and the protected RNA fragments were separated on a 5% denaturing polyacrylamide gel. The resultant products were assessed by autoradiography using BioMax MS film at −70 C.

Cell culture

ROS17/2.8 osteoblasts were grown in a 1:1 mixture of DMEM and Nutrient Mixture F-12 (Life Technologies, Inc.), as described previously (18). The MC3T3-E1 osteoblast cell line was grown in α-MEM (Life Technologies, Inc.) and was grown in α-MEM supplemented with 0.13 mM ascorbic acid and 5 mM β-glycerol phosphate for differentiation, as described previously (18). COS-7 cells were maintained in DMEM (Life Technologies, Inc.) (19). C5.18 chondrocytes (20) were provided by Dr. Jane E. Aubin (University of Toronto, Toronto, Canada) and maintained in α-MEM containing 15% FBS. All cell lines, except C5.18, were supplemented with 10% (vol/vol) FBS. All cultures were supplemented with 100 μg/ml penicillin and streptomycin and cultured in a humidified incubator with 5% CO_2 at a temperature of 37 C.

Preparation of constructs

To generate the *Phex* reporter gene constructs, 2790 bp of the genomic fragment from −2736 to +54 of *Phex* was amplified with PCR SuperMix High Fidelity (Life Technologies, Inc.) using a set of primers (−2736 F, 5′-GGGGTACCGCCAGTGGGGTCTTGTATGT; +54 R, 5′-GGGGTACCAGATTTCTGCTATGACAGCC), and the resultant product was subcloned into *Kpn*I site of pGL2-Basic vector named p2736*Phex*-luc. Sequentially, 5′-deletion with different primers generated constructs from −2736 to −22, including p1606*Phex*-luc with primer −1606 F, 5′-GGGGTACCATGCATTTGCTGTCACATAT; p964Phex-luc with primer −964 F, 5′-GGGGTACCTGGTTAAGATATGTTAGG; p472*Phex*-luc with primer −472 F, 5′-GGGGTACCCTTAATCCTCAGGAAGCT; p178*Phex*-luc with primer −178 F, 5′-GGGGTACCAGTTCCAGTCCAAACCATCA; p130*Phex*-luc with primer −130 F, 5′- GGGGTACCTTGCACTGCAATGGACTATG; and p22 F with primer −22 F, 5′-GGGGTACCGGGACTAACACACTGAAAGAGT. We used the previously described SOX9 (21) and the *Cbfa1* mammalian expression constructs (22).

Transient transfection and reporter assays

Transient transfection experiments were performed using the TransFast Reagent (Promega Corp., Madison, WI) as described previously (22). Briefly, cells were plated at a density of 1.5×10^5 cells/well in 6-well plates 16 h before transfection. We used 1.0 μg *Phex* promoter constructs and 0.25 μg pSV β-galactosidase, and in cotransfection experiments we added 0.5 μg SOX9 or *Cbfa1* expression constructs. Luciferase activity was measured using the Luciferase Assay Kit (Promega Corp.). β-Galactosidase activity was measured by β-Galactosidase Enzyme Assay System (Promega Corp.). For the transient transfection studies, the luciferase activity was normalized by β-galactosidase activity by dividing luciferase activity by β-gal. The relative luciferase activity was then calculated by dividing the normalized luciferase activity by that obtained with the empty pGL2-Basic vector. We observed higher β-gal activity in Cos-7 cells, but this did not influence the relative comparisons between cell lines. For stable transfection studies in MC3T3-E1 osteoblasts, the luciferase activity was normalized for cell number by dividing luciferase activity by DNA content. The relative luciferase activity was then calculated by dividing the normalized luciferase activity by that obtained with the MC3T3-E cells stably transfected with the empty pGL2-Basic vector. For the stimulation studies, cells were incubated in serum-free medium containing 0.1%BSA for 24 h before adding different concentration of various reagents, including PTH 1–34, forskolin, 12-*O*-tetradecanoylphorbol-13-acetate (TPA; Sigma, St. Louis, MO), and $1,25(OH)_2D_3$, as described previously (17). We previously have determined the concentration of PTH and forskolin to stimulate cAMP and the concentration of TPA to stimulate PKC activity in osteoblasts (data not shown).

Stable transfection of MC3T3-E1 with Phex promoter/luciferase reporter construct

Stable transfection of MC3T3-E1 was performed by a pooled protocol as described previously (18). MC3T3 cells were cotransfected with p2736Phex-luc and pSV2-neo in a 15:1 molecular ratio. Transfectants were selected in the presence of 700 μg/ml G418 for 14 d. The cells were plated in α-MEM complemented with ascorbic acid and β-glycerol phosphate. The luciferase activities were measured on different days using the Luciferase Assay Kit (Promega Corp.). Total DNA content was determined using Picogreen dsDNA Quantitation Reagent and Kits (Molecular Probes, Inc., Eugene, OR).

Statistical analysis

We evaluated differences between groups by one-way ANOVA. All values are expressed as mean ± SEM. All computations were performed using the Statgraphic statistical graphics system (STSC, Inc., Rockville, MD).

Results

Isolation and cloning of the mouse Phex promoter and its 5′ untranslated region (UTR)

At the time we initiated our studies the *Phex* promoter had not been isolated or cloned and the full extent of the 5′ UTR had not been defined. To locate the *Phex* promoter, we isolated and cloned the region of the mouse *Phex* gene containing exon 1 and upstream sequences (Fig. 1). RT-PCR and RNase protection analysis (Fig. 2) demonstrated that exon 1 contains 556 bp of the 5′ UTR and evidence for TATA-containing promoter in the 2736 bp of upstream sequence (Fig. 1). Using computer prediction programs, we initially located the putative transcription start site 556 bp upstream of the ATG in exon 1. We performed RNase protection to confirm the location of the transcription start site (Fig. 2A). Using a riboprobe (−130, +132), which contained 262 oligonucleotides overlapping the putative start site, we identified a single 132 bp protected fragment in both mature MC3T3-E1 and TM-Ob osteoblasts. To provide additional evidence for the location of the transcription start site, we performed RT-PCR using intron-spanning primers flanking the initiation sites and RNA derived from differentiated, *Phex*-expressing MC3T3-E1 osteoblasts (Fig. 2B). Forward primer +1 F in combination with the reverse primer +697 R in exon 2 generated the predicted size product. In contrast, the upstream primer −22 F in combination with +697 R did not amplify any product, indicating that transcription initiation begins in the segment flanked by primers −22 F and +1 F (Fig. 2B). Consistent with the designated start site, we found a consensus TATA-box at position −32 to −23 from the transcription start site (Fig. 1).

We used the mouse *Phex* promoter sequence to search for homologies in the corresponding 5′-flanking region of the human *Phex* gene that has recently been sequenced (16). The overall homology was 71% over the 2736-bp region corresponding to the mouse *Phex* promoter. In addition, we found five nearly identical regions between the 2736 bp of the mouse *Phex* promoter and the 5′-flanking region of the human PHEX gene. These include an identity of 96% over the first 182 bp from the transcription start site (position −182 to +1) and regions exhibiting more than 82% identity at positions −2425 to −2233 (192 bp), −1959 to −1906 (53 bp), −1317 to −1267 (50 bp), and −541 to −491 (37 bp) of the mouse promoter. In addition, we observed in both human and mouse sequence many conserved consensus elements including a GRE at position −2634 (23); several SRY sites at positions −2340, −2199, −700, −468, −388, and −313 (24); GATA-1 at positions −2402, −1920, and −166 (25); AP1 at position −673 (26); MZF1 at position −1909 (27); HMG box at position −1900 (28); E-box at position −1934 (29); EF1 repressor at position −1643; Bcd at the position −1573 (30); AP-4 at position −1263 (31); NF-Y at position −1057 (32); C/EBP at position −928 (33); C-Rel at position −897 (34); Lmo2 complex at position −625 (35); Nkx-2.5 at position −143 (36); Skn-1 at position −94 (37); GATA-2 at position −58 (38); and a TATA box at position −32 (39). We also observed several putative *cis*-acting elements in the mouse sequence that were not conserved with the corresponding human 5′-flanking region, including a possible AML-1a binding site at position −277 (40).

Cell specificity and abundance of Phex mRNA

Before proceeding with the functional analysis of the 5′-flanking region of the *Phex* gene, we identified *Phex*-expressing and -nonexpressing cells. By RNase protection analysis we confirmed the osteoblast expression and differentiation stage dependency of *Phex* expression (Fig. 2A and Fig. 3). We detected a high level of *Phex* expression by RNase protection assay in mature MC3T3-E1(14 d of culture) and TM-Ob (15 d of culture) osteoblasts, but failed to detect *Phex* transcripts in NIH3T3 fibroblasts and undifferentiated TM-Ob cells cultured for 5 d. Using RT-PCR analysis, we detected *Phex*. transcripts in differentiated MC3T3-E1 osteoblasts, C5.18 chondrocytes and ROS17/2.8 osteosarcoma cells, but not in COS-7 cells (Fig. 3).

Functional characterization of the Phex promoter in ROS17/2.8 osteoblasts

To begin to examine the transcription of the *Phex* gene, we evaluated serial deletions of *Phex*-luc promoter constructs in ROS17/2.8 osteosarcoma cells (Fig. 4). Progressive 5′-deletion mutations of the promoter revealed a bimodal pattern of functional activity in transiently transfected ROS17/2.8 cells (Fig. 4). The full-length promoter construct (−2736/+54) was active in ROS 17/2.8 cells, and displayed a progressive reduction in promoter activity as the 5′-deletion mutations approached 1606 bp, relative to the transcription start site. Further deletions from 1606 to 946 bp, however, lead to an increase in promoter activity that exceed that of the full-length promoter. Successive deletions from −946 bp lead to a progressive reduction in promoter activity in ROS17/2.8 cells, resulting in a near complete ablation with −22/+54 construct.

Cell-restricted activity of the 2.7 kb and 946 bp Phex promoter constructs

Phex transgene constructs containing 946 bp and 2736 bp of upstream sequence (p946*Phex*-luc and p2736*Phex*-luc) were further tested to define the DNA regulatory regions that might direct the tissue-specific expression of *Phex* (Fig. 5). Because *Phex* is strongly expressed in osteoblasts and chon-

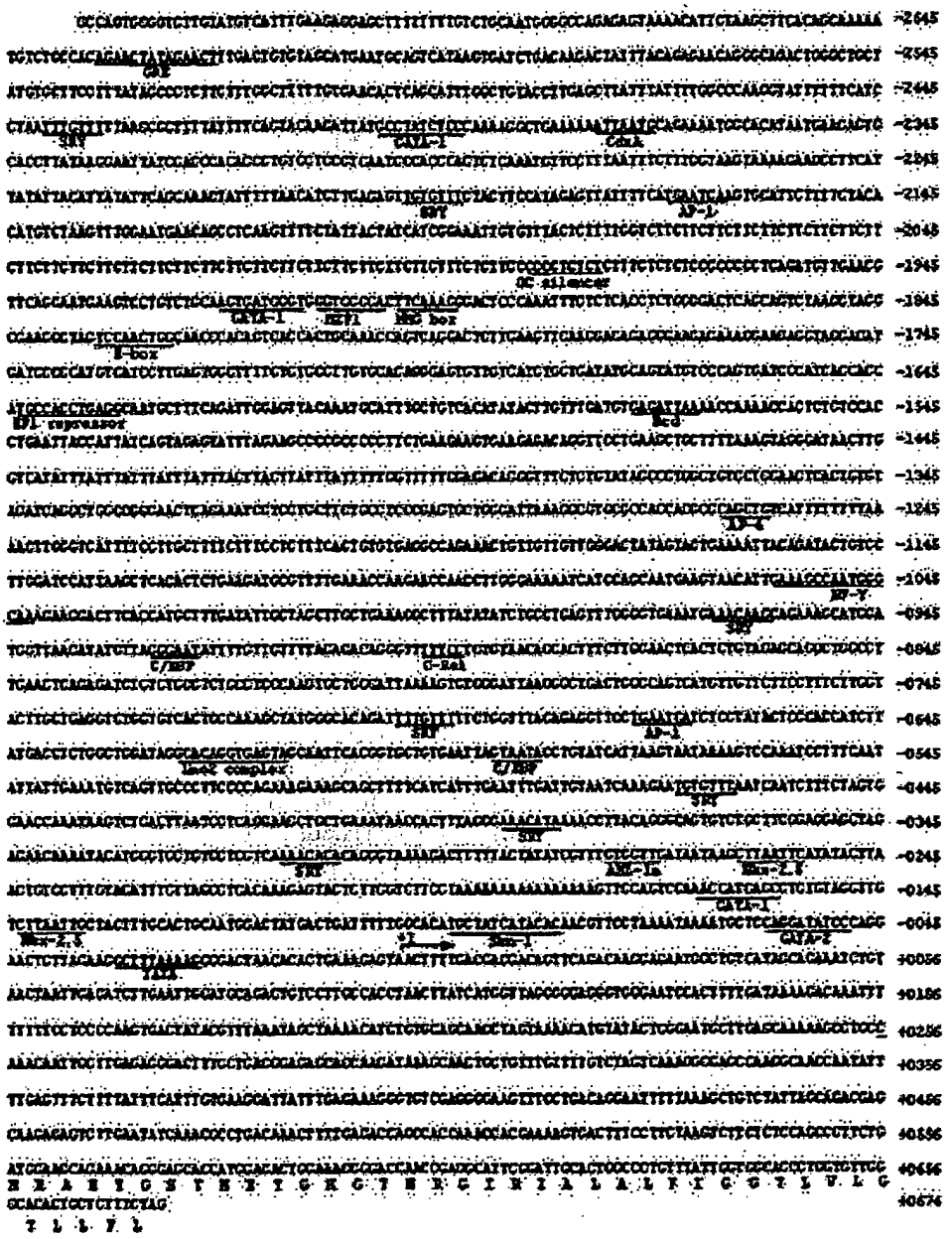


FIG. 1. Nucleotide sequence of the 5′-flanking region of the *Phex* gene. The sequence encompasses 3410 bp of the 5′-flanking region of the *Phex* gene. The transcription start site based on RNase protection and RT-PCR analysis (Fig. 2) is shown by an *arrow*. Consensus *cis*-acting elements for known transcription factors conserved between mouse and human are *underlined*. The sequence is *numbered* relative to the transcription start site.

drocytes but is not expressed in COS-7 cells, these two *Phex* promoter constructs were transiently transfected into these three different cell lines. The full-length promoter construct (−2763/+54) was expressed to a higher extent (approximately 3-fold) in C5.18 chondrocytes compared with COS-7 cells. The expression was also significantly greater in ROS 17/2.8 osteosarcoma cells compared with COS-7 cells, but was quantitatively less than expected from the level of Phex mRNA in osteoblasts (Fig. 3). In all three cell lines, the overall activity of the (2736/+54) construct was lower than that observed with the truncated promoter construct (−946/+54). Indeed, deletion of the initial 1789 bp of the promoter resulted

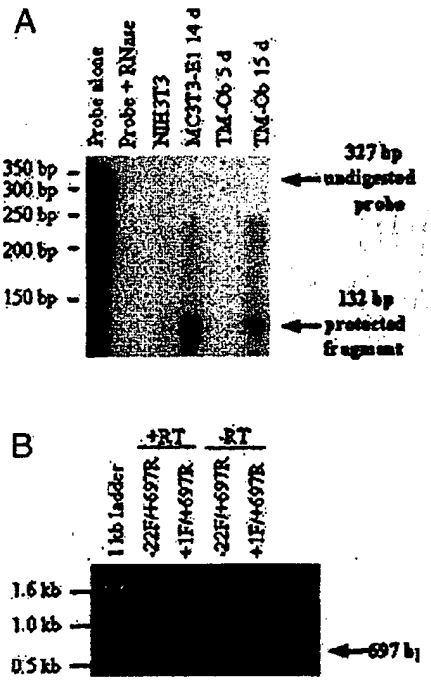


FIG. 2. Determination of the transcription start site of the mouse *Phex* gene. A, RNase protection was performed with 20 μg total RNA hybridized to ^{32}P-labeled 327-bp riboprobe complementary to nucleotide from −130 to +132. The approximate position of the 5′ ends was estimated by the size of the protected fragment. B, RT-PCR mapping was performed with total RNA derived from mature MC3T3-E1 osteoblasts. The reverse primer, *Phex*+697R in exon 2 was used in combination with respective forward primers *Phex* −22F, *Phex* +1F. The position of the forward primers relative to the transcription start site is shown in Fig. 1. +RT, with RT; −RT, without RT.

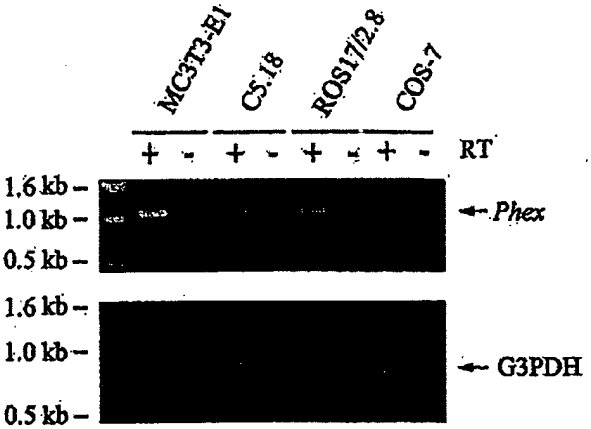


FIG. 3. RT-PCR analysis of *Phex* mRNA expression in various cell lines. RT-PCR was performed with 2.5 μg total RNA from MC3T3-E1, C5.18, ROS 17/2.8, and COS-7 cells. MC3T3-E1 cells were induced to differentiate by 14 d of culture in α-MEM containing ascorbic acid and β-glycerol phosphate. *Phex* was amplified by PCR using the forward primer +1866F and the reverse primer +2943R (*top*). A 1.1-kb product was amplified from 14-d-old MC3T3-E1, C5.18, and ROS17/2.8 osteoblasts. No product was amplified from COS-7 cells. G3PDH primers, which amplify a 0.9-kb product, were used as controls for RNA integrity (*bottom*). +RT, with RT; −RT, without RT.

in nearly a 3-fold increase in promoter activity in C5.18 chondrocytes and nearly a 2-fold activity in ROS17/2.8 osteosarcoma cells and COS-7 cells. These findings suggest that the distal region of the promoter from −946 to −2736 contains inhibitory elements and the proximal 946 bp of the promoter has regulatory elements that contribute to the differential expression of *Phex* between osteoblasts, chondrocytes and non-osteoblastic cell lines. The results in COS-7 cells need to be interpreted with caution, because they may not simply reflect the nonosteoblastic phenotype of the cells, but the effects of their expression of SV40 T antigen.

SOX9 enhances Phex promoter activity

Because of the high levels of expression of chondrocytes, we tested the effect of transcription factors regulating chondrogenesis on *Phex* promoter activity. SOX9 is a transcription factor that plays a key role in chondrogenesis (21). To examine whether SOX9 regulates the transcriptional activity of the *Phex* gene, the full-length *Phex* promoter construct (−2736/+54) was cotransfected with a SOX9 expression vector into ROS17/2.8 cells(Fig. 6A). SOX9 overexpression in C5.18 chondrocytes, ROS17/2.8 osteosarcoma cells, and COS-7 cells enhanced by approximately 6-, 4-, and 3-fold, respectively, the transcriptional activity of the full-length *Phex* promoter. In contrast, SOX9 failed to stimulate activity above the control plasmid using the (−946/+54) construct (Fig. 6B). These observations suggest that SOX9 may regulate *Phex* gene transcription through upstream SRY-like elements, possibly at positions −2340, and −2199 (23). Additional studies will be needed to confirm the importance of the putative SRY sites and identify the transcription factors that bind to them. Other factors also are likely to be involved in mediating the preferential expression of *Phex* in chondrocytes. Indeed, SOX9 mRNA was detected in all three cell lines (Fig. 6C), but its ability to up-regulate promoter activity was greatest in chondrocytes. In addition, the unexpected high activity of the *Phex* promoter in COS-7 cells may be explained by their high level of endogenous SOX9 transcripts.

Failure of other factors to regulate the Phex promoter

We performed additional studies evaluating the effects of *Cbfa1*, PTH, calcium, forskolin, 1,25$(OH)_2D_3$, phosphorus, and TPA on either *Phex* promoter construct (−2736/+54) in ROS17/2.8 cells. In contrast to the potent stimulation by SOX9 (Fig. 6), none of these factors stimulated *Phex* promoter activity (data not shown). In particular, we observed no stimulation by cotransfecting the *Cbfa1* expression plasmid (22) with the *Phex* promoter, whereas this *Cbfa1* expression plasmid has been previously used by us to stimulate an osteocalcin promoter construct (22). Similarly, 1,25$(OH)_2D_3$ at concentrations that stimulated an osteocalcin promoter construct failed to stimulate the *Phex* promoter construct in ROS17/2.8 (Fig. 7) or MC3T3-E1 osteoblasts. PTH at concentration ranging from 5–100 nM as well as the addition of calcium to the media at concentrations ranging from 1–5 mM also failed to stimulate *Phex* promoter activity in ROS17/2.8

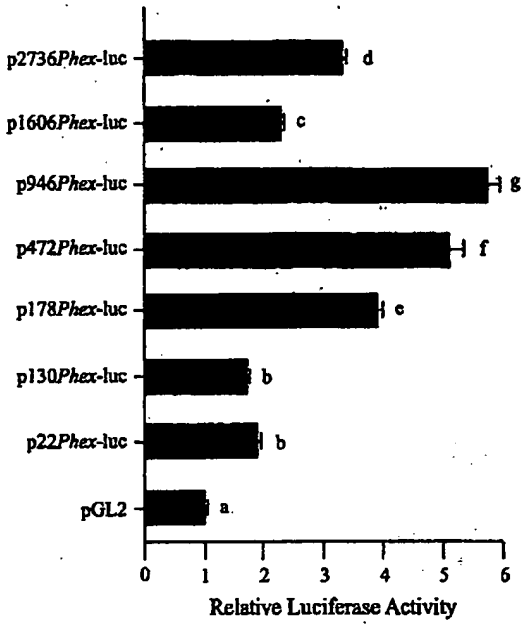


FIG. 4. Function analysis of 5′ deletion mutants of the mouse *Phex* promoter-luciferase chimeric gene by DNA transient transfection experiments. Activity of deletion mutants of the *Phex*-luc constructs were assessed in ROS 17/2.8 cells that express the *Phex* (Fig. 3). The promoter/reporter construct, p2736*Phex*-luc, consists of sequence −2736 to +54 subcloned into the pGL2-Basic vector. Deletions were generated by PCR as described in *Materials and Methods*. In DNA transfection experiments, successive deletions of *Phex*-luc from −2336 to −472 decreased the level of expression 40%, whereas deletions from −130 to −22 abolished activity, indicating critical *cis*-acting elements are present between −947 and −472, as well as between −130 and −22. Luciferase activity is relative to pGL2-Basic vector alone and expressed as a ratio to β-galactosidase activity to correct for transfection efficiency. Values represent the mean ± SEM of a minimum of three separate transfection experiments. Values sharing the same *superscript* are not significantly different at $P < 0.05$.

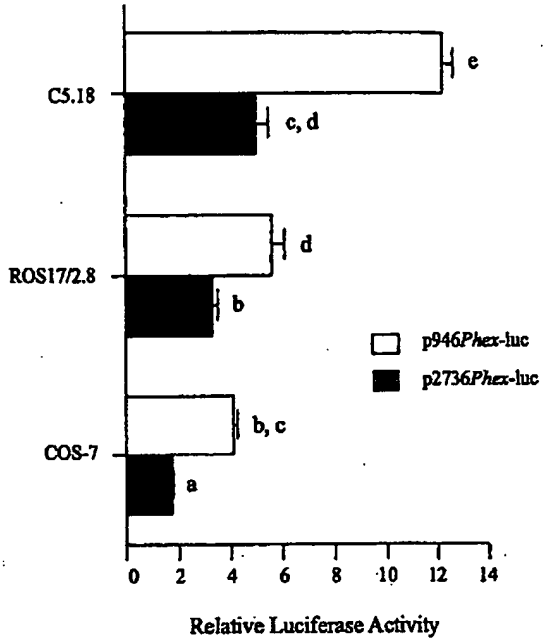


FIG. 5. Function analysis of the murine *Phex* promoter in different cell lines: evidence for positive and negative regulators. We compared constructs p2736*Phex*-luc and p946*Phex*-luc that, respectively, contain 2790 bp and 946 bp of the 5′-flanking portion of the *Phex* gene inserted immediately upstream from the firefly luciferase in pGL2-Basic. One microgram of p*Phex*-luc and 0.25 μg pSV β-galactosidase were cotransfected into cells using TransFast reagent. The luciferase activity was normalized by β-galactosidase activity and expressed relative to the pGL2-basic vector. The p946*Phex*-luc construct demonstrated greater activity in C5.18 chondrocytes and ROS17/2.8 osteoblasts compared with non-*Phex*-expressing COS-7 cells. This suggests the presence of elements in the *Phex* promoter directing cell-restricted expression as well as nonspecific elements directing high-expression in COS-7 cells. In contrast, the p2736*Phex*-luc construct expressed less activity in all cell types, but maintained the relative greater activity in *Phex*-expressing chondrocytes compared with osteoblasts and COS-7 cells. The values represent the mean ± SEM of three separate transfections. Values sharing the same *superscript* are not significantly different at $P < 0.05$.

cells (data not shown). Finally, neither stimulation of cAMP by forskolin at a dose ranging from 10^{-7} to 10^{-5} M, or activation of PKC by treatment of cells with TPA at concentrations of 0.16 μM and 1.6 μM stimulated the *Phex* promoter construct in MC3T3-E1 osteoblasts (data not shown).

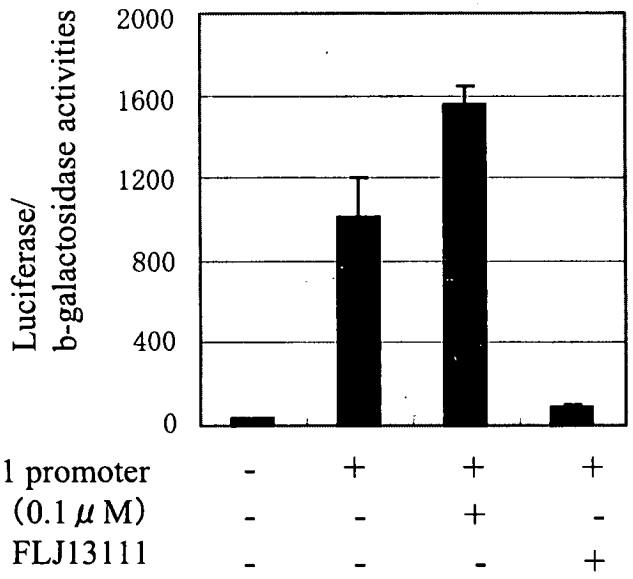
Developmental Stage-specific Transcriptional up-regulation of the Phex promoter in MC3T3-E1 cells

Different stages of maturation were accomplished by growing MC3T3-E1 cells in media containing serum with supplemented with ascorbic acid and β-glycerol phosphate for various time periods (18). To confirm the maturational up-regulation of *Phex* expression, we evaluated *Phex* promoter activity in stably transfected MC3T3-E1 osteoblasts after 4 and 9 d of culture (Fig. 8A). We found that *Phex* promoter activity relative to the empty vector alone increased during this maturation period, paralleling the culture duration dependent up-regulation of endogenous *Phex* mRNA in these cells (Fig. 8B).

Discussion

To initiate studies directed at identifying pathways regulating *Phex* expression, we have isolated, sequenced, and characterized the functional activity of the murine *Phex* promoter in chondrocytes, osteoblasts, and COS-7 cells. A single transcription start site was mapped, and a promoter construct containing up to 2736 bp of the 5′-flanking sequence was identified and shown to display functional activity in cells expressing *Phex* endogenously. The promoter region that we identified is the same as that independently isolated from a mouse genomic BAC clone (41). Searching the DNA database with the mouse promoter sequence identified a corresponding highly homologous region upstream of the human PHEX coding sequence and conservation of many of the same putative regulatory *cis*-acting elements in the mouse and human 5′-flanking regions (Fig. 1). We have begun investigations with the murine *Phex* promoter to de-

Fig.9′


Luciferase/
b-galactosidase activities
2000
1600
1200
800
400
0
3111 promoter - + + +
one (0.1 μM) - - + -
FLJ13111 - - - +